Predictors of Phonological Awareness in Five-Year-Olds: Contrasting Measures of Phonological Knowledge

A THESIS
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Challie Johnson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF ARTS

Dr. Benjamin Munson

May 2022

Acknowledgements

The warmest of thanks to Mx. Munson, a deep well of knowledge and zeal. Thank you for the privilege of stewarding an excerpt of the work you began decades ago. You stir others towards better questions and thorough pursuit.

Thank you, Michelle Erskine, Jan Edwards, and Meg Cychosz, for paving the way in data collection and inquiry. This analysis, both whole and in part, belongs to you.

Thank you, Dr. Kerry Ebert, Dr. Panayiota Kendeou, for offering your time and energy to further this project and uphold the standards of research.

Thank you, Cammie McBride, Tiffany Hogan, for collaboration and inspiration.

Abstract

A host of variables have been proposed as correlates of phonological awareness (PA). For young children, there are strong theoretical and experimental arguments that PA is predicted by different components of phonological knowledge, speech perception, nonword repetition accuracy and speech production, as well as vocabulary and nonverbal cognition; however, research has yet to consider these variables together within a coherent model. The current study seeks to extend the developmental framework of PA with a population of children at the approximate age of school entry. Data from 92 typically developing children (age 4;4-5;4) were selected from a longitudinal study on vocabulary and PA based on completion of tasks of interest to the current study. Performance was modeled using path analysis to evaluate a variety of models of PA, including models inspired by Rvachew and Grawburg (2006), and additional models including a measure of higher-level phonological knowledge, the lexicality effect (Cychosz et al., 2021), and nonverbal cognition. Results challenged previous models that support speech perception as a strong predictor of PA and show speech production, instead, as a robust predictor of phonological knowledge. Variance in PA was enhanced by the addition of a lexicality effect but not nonverbal cognition.

Table of Contents

List of Tables

List of Figures

Phonological awareness (PA) represents the ability to access the underlying sound structure of a word, otherwise known as its phonological representation, and to identify and manipulate its phonemes. PA is a fundamental pre-literacy skill and stands as one of the most powerful predictors of reading delays and disabilities in children. That is, poor quality lexical representations are a well-known presage of reading difficulty. While knowledge of reading development continues to grow, the role of PA in the development of accurate and efficient decoding remains evident (Petscher et al., 2020).

For preschool children, measures of PA, in combination with measures of letter knowledge, predict up to 54% of one important component of reading ability, decoding skills, which refers to the ability to apply letter-sound relationships to read written words (Lonigan et al., 2000). Furthermore, measures of phonological representations stand as one of three statistically significant predictors of dyslexia, along with letter naming, and phoneme identification (Elbro et al., 1998). One cannot underestimate the potential cascade of impacts experienced secondary to reading difficulties nor, therefore, the importance of earliest and most effective identification and prevention.

When a lack of early PA sounds the *casus belli*, research shows that intervention to remediate PA skills effectively answers the call for young children at risk of reading difficulties. The 2000 Report of the National Reading Panel detailed a comprehensive meta-analysis pertaining to the effects of direct instruction of PA; regardless of age or instructional format, significant gains in PA, reading, and spelling were observed (National Institute of Child Health and Human Development [NICHD], 2000). For this reason, evidence-based methods of reading instruction continue to directly support PA skills (Petscher et al., 2020). The efficacy of PA intervention spurs teachers and clinicians to begin PA intervention with urgency with the near 10% of unexpectedly poor and dyslexic school-age readers falling behind (Wagner et al., 2020).

However, what if children with reading delays could be identified and provided with focused PA support prior to such a critical moment in reading development? Research across clinical and educational disciplines continues to search for ways to improve reading outcomes prior to school entry.

A bottom-up perspective of PA, that taken in the current study, offers the potential for prevention of reading difficulties through its early developing skills. While

phonological processing abilities in young children can be formally measured beginning at the age of four, fundamental contributors of PA, including speech perception, receptive and expressive vocabulary, speech production, nonverbal cognition, and others may be critical predecessors. These components proceed from birth, develop measurably prior to age four, and continue to progress alongside explicit PA knowledge throughout early to late childhood. Therefore, understanding the relationships between the fundamental elements contributing to PA may be the key to developing a theoretical framework for both the development of PA and prevention of reading-related disabilities long before literacy instruction begins.

The current study will first discuss previous research examining associations between speech perception, receptive and expressive vocabulary, speech production, nonword repetition accuracy, and nonverbal cognition with measures of PA. While each of these variables have complex and interconnected relationships with one another, the focus of this review remains on their respective correlations with PA. For example, a large body of research has examined how speech perception and production relate to one another. These correlations remain outside of the scope of the current investigation, and thus will not be discussed here. Then, several models outlining concurrent and predictive correlations across variables of PA will be examined.

Phonological awareness

Phonological awareness describes an individual's ability to consciously identify distinct sounds and sequences within a speech stream (Anthony & Francis, 2005). Awareness of whole word structures emerges first while awareness of increasingly smaller units within words (i.e., rime, syllable, and individual phonemes) develops with continued exposure. Lonigan et al. (2000) demonstrated the continuum of PA development in a longitudinal study of 2 cohorts of children from early to late preschool. At Time 1, a younger sample ranged from 25-61 months and the older 48-64 months. Time 2 testing was completed 18 months later for the younger cohort and 12 months later for the older. Phonological sensitivity was measured at both time points by onset-rhyme (e.g. Which of these words does not rhyme: *pail, mail, coat*), alliteration identification (e.g. Which of these words does not start with the same sound: *cat, sun, car*), blending

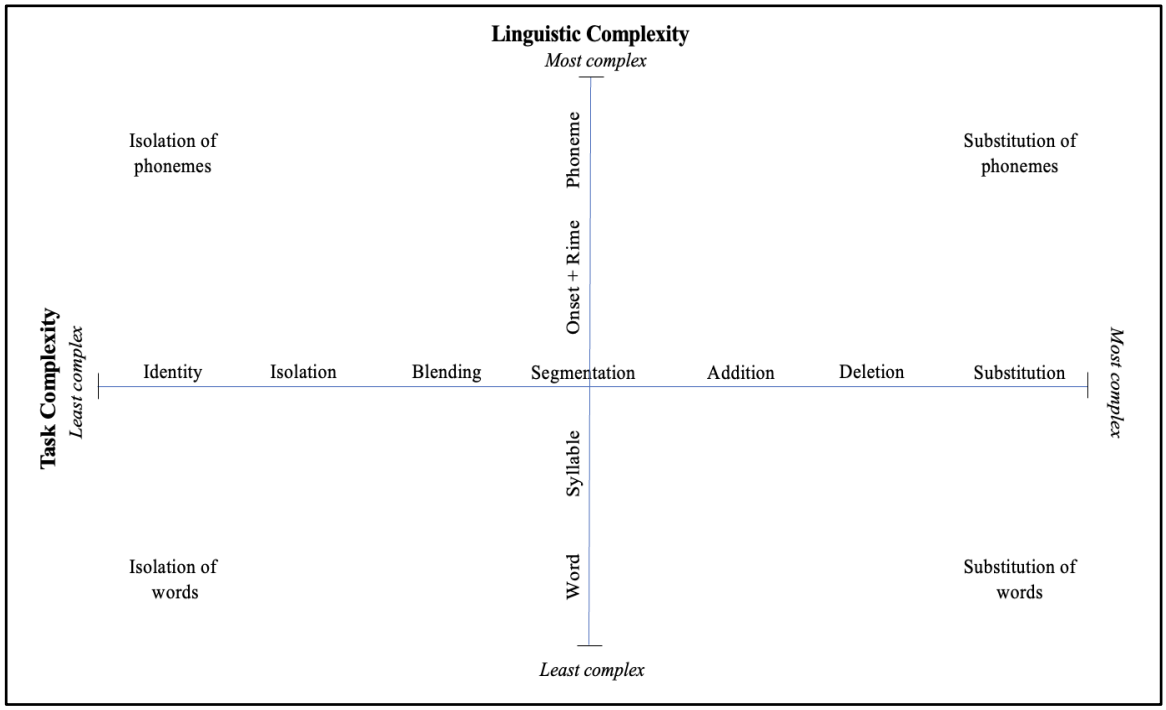
(e.g. What do you get when you say *bat… man* together?) , and elision (e.g. What do you get when you say *batman* without *man?*) tasks. For both cohorts, phonological sensitivity measures at Time 1 were significantly predictive of the same measures at Time 2. Results demonstrated that measures of phonological sensitivity represent a continuum of development rather than distinct abilities, with multiple skills advancing and refining simultaneously with experience and instruction. This progression of awareness is supported by general linguistic experience as well as a range of related linguistic and cognitive factors, as will be discussed in the following sections.

Intuitively, then, tasks to measure PA are arranged hierarchically based upon the explicit complexity of the task, the linguistic complexity of the presented stimuli, or both. Linguistic complexity refers to many things. In the context of PA tasks, these include the real or nonword status of the stimuli as well as the probability of the targeted sound sequence in the listener's native language. Regarding task, the ability to discriminate between individual linguistic units precedes the ability to identify or compare them. Broadly, manipulation, deletion, blending, and segmentation skills then proceed in turn. These generalizations are supported by McBride-Chang (1998), who sought to determine if specific manipulations within PA tasks influence performance more readily than others. One hundred thirty-six children were presented with three widely accepted PA tasks: phoneme deletion (e.g., Say *cat* without */k/*), position analysis (e.g., What sound do you hear at the end of *cat*), and phoneme segmentation (e.g., Say the sounds in *run)*. Children's performance supported the primary hypothesis: each of the three manipulations varied in difficulty, with phoneme segmentation being the most difficult task.

In this sense, at the phoneme level, segmentation tasks will prove much more difficult for younger children whose PA skills may be better estimated through phoneme deletion or position analysis tasks (McBride-Chang, 1995). See Figure 1 for one interpretation of complexity in relation to measures of PA. Nevertheless, this collection of skills, varying by type of task as well as sequence or unit of sound, have been accepted as valid measures of PA and used to analyze variables both predictive *of* and predicted *by* PA.

Prior to 1998, theories of PA proposed only three underlying component skills: speech perception, short term memory (STM), and general cognition (McBride-Chang, 1995). Speech perception represented the ability to discriminate the auditory characteristics of stimuli and perform the correct manipulations. Short-term memory allowed retention of the stimuli throughout the task while general cognitive ability was needed to manipulate the stimuli accordingly. In recent decades, these predictive components have proven to be robust predictors of PA both individually and in various frameworks. More recently, speech production, expressive and receptive language, and nonword repetition skills have been proposed as important predictors of PA.

Figure 1. PA Task Complexity, inspired by McBride-Chang (1995)



Nonverbal cognition

Logic asserts that all PA tasks demand an underlying level of cognitive ability. One must thoughtfully consider the task at hand, act upon the given stimuli, and execute fully. Across the literature, IQ similarly represents a broader measure of an individual's ability to cognitively manipulate stimuli. No comprehensive theoretical model has proposed direct effects of cognitive level on performance on PA tasks; however, its role has been considered in a handful of correlational analyses.

McBride-Chang (1995), as discussed above, proposed that IQ, along with perception and short-term memory, are essential components of PA. This hypothesis aligned with previous correlations between cognitive ability and performance on phonological processing tasks. Third- and fourth- grade children completed four subtests on the Wechsler Intelligence Scale for Children – Third Edition (WISC-III; Wechsler, 1974) and three tests of PA. Results demonstrated that IQ had moderate to high contributions to the PA construct. That is, children with higher nonverbal cognition generally performed better on the PA tasks.

With aims to assess phonological working memory prior to age 4, Gathercole and Adams (1993) presented children ages 2-3 with phonological working memory tasks, including digit span, word and nonword repetition, as well as standardized measures of articulation and receptive vocabulary. Nonverbal intelligence, as measured by the McCarthy Scales of Children's Abilities (MSCA; McCarthy, 1970), demonstrated only nominal correlations with all principal measures (Gathercole & Adams, 1993).

Meanwhile, Elbro et al. (1998) considered the impact of cognition on predicting dyslexia in grade 2. Given variability across PA tasks, perhaps the role of cognition could be viewed when considering other, closely related measures of PA and reading success. Though three significant predictors of dyslexia were identified, namely letter and phoneme identification and distinctness of phonological representations, there were no main effects of cognitive measures in predicting later reading development (Elbro et al., 1998).

Even so, measures of nonverbal intelligence, if not addressed, may present as a confounding variable in the current models. Despite being a nonverbal measure, by definition, IQ may covary with general language ability. In Dollaghan and Campbell

(1998), children aged 6-9 with language impairment and their language-normal counterparts were compared on measures of nonword repetition, receptive and expressive language, and nonverbal intelligence. A significant group difference in performance on the Test of Nonverbal Intelligence-Revised (TONI-R, Brown et al., 1990) was revealed. Thus, suggesting that children with language impairment demonstrate lower cognitive skills than their typically developing peers, despite overall performance within normal limits (Dollaghan & Campbell, 1998).

In all, modest variations observed across principal measures of language and phonological memory support nonverbal intelligence as a standard battery for tasks, like oral PA tasks, requiring any form of cognitive manipulation.

Speech perception

Oral PA awareness tasks require speech perception to accurately perceive, store, and categorize presented stimuli. Therefore, one possible explanation for impaired PA is difficulty encoding the speech signal during perception. The ability to perceptually distinguish increasingly small units precedes and bolsters the development of early phonological representations. An individual with impaired perception of a stimulus will likely have loosely categorized or incomplete representations of the stimulus in their working memory. It is difficult to imagine how intact PA could develop in an individual whose perception is persistently poor. Even with PA intact, a lack of adequate speech perception would cause significant hindrance in analyzing or manipulating the linguistic unit.

Task and linguistic complexity are important considerations for measures of speech perception. Tasks employed across the literature may include identification (e.g., identify stimuli X as either */r/* or */w/)*, discrimination (e.g., describe stimuli X and Y as same or different), or ratings of similarity at the phoneme, syllable, or word level. Each task presents unique demands on the listener based upon age and acuity.

For children, prolonged exposure to spoken language provides one form of practice in perceiving acoustic units within connected speech. Similarly, perception is particularly supported by growth in vocabulary skills, in that sensitivity to underlying structure heightens in familiar words. For example, a child familiar with the words *abject*

and *conjure* is more likely to accurately perceive the word *abjure* when first introduced than a child who does not have these similarly structured words established in their lexicon. The known words support both the ability to correctly perceive and create a phonological representation of the novel word. In this way, language-dense stimuli may not manifest a clear view of an individual's basic perceptual skills and instead reflect linguistic knowledge (Sénéchal et al., 2004). Traditional speech perception tasks at the syllable or phoneme level seek to reduce the impacts of language and memory demands.

Of note, speech perception tasks can also be useful in examining how different characteristics of words influence performance. Several studies support a unique and significant variance in PA from spoken word recognition (Garlock et al., 2001; Walley et al., 2003). Specifically, research has examined how variations in age of acquisition and neighborhood density of the word-recognition stimuli affect word-recognition performance, and how those influences might relate to PA. Much of this work has evaluated the *lexical restructuring hypothesis* (LRH; Metsala & Walley, 1998), which posits that children's early representations of words are holistic. Restructuring, the process of accessing the sublexical units through gradual segmentation, is mediated by the same factors that comprise PA. LRH proposes that spoken word recognition, specifically the ability to recognize words from sparse neighborhoods, is the strongest indicator as to the extent of restructuring (Metsala & Walley, 1998).

Strong correlations between speech perception and PA were first reported in McBride-Chang's (1998) cross-sectional study of the extent to which different underlying factors influence components of PA in typically developing children. One hundred and thirty-six students in grades 3-4 completed three computer-based forced-choice identification tasks using three speech continua: */ba/-/wa/* (manipulating the duration of the formant transition), *slit-split* (manipulating the temporal gap between the /s/ and the /l/) and *bath-path* (manipulating the voice onset time of the initial bilabial stop). Through structural equation modeling these three measures composed a single latent variable representing speech perception. Likewise, the PA latent variable was based on performance on three common PA tasks: phoneme deletion, position analysis, and phoneme segmentation. Regression analysis showed that speech perception predicted moderate variance to the PA construct. This was true even when the influence of general

cognitive ability and short-term memory were statistically controlled; that is, speech perception predicted a unique portion of variance in PA (McBride-Chang, 1998).

Further research suggests that a one-way variance may not fully represent the strength of this correlation. Predictive relationships between speech perception, PA, and articulation were examined by Sénéchal et al. (2004) in children 48-66 months. At the word level, results demonstrated a phoneme awareness task, instead, was a stronger predictor of performance on a word discrimination speech perception task than the reverse, leading the question of whether the relationship between these variables could be interpreted bidirectionally. Regardless of order, a correlation between speech perception and PA was evident across regression analyses (Sénéchal et al., 2004).

Still, recent evidence has questioned the strength of correlation between speech perception and PA. Results from Erskine et al. (2020) demonstrated the predictive performance of several component variables, including a minimal pair discrimination task, with children ages 2-3. Specifically, researchers compared performance on a minimal pair discrimination task to performance on standardized measures of PA two years later. Surprisingly, children's speech perception ability did not yield a significant effect (Erskine et al., 2020). The authors conceded to the findings of McBride-Chang (1995), stating that mixed results are expected given the differential measures of speech perception employed. Nevertheless, the relationship between speech perception and PA continues to receive both rational and theoretical backing.

Various analyses have considered correlations between PA and speech perception specific to children with speech sound disorders (SSD). SSD refers to extremely inaccurate speech in the absence of an obvious sensory, motor, structural, or cognitive problem that would affect speech sound development. Deficits in articulation are hypothesized to be a direct reflection and reinforcement of inaccurate perception. In this way, speech production errors can hinder the development of accurate phonological representations and therefore result in deficits in PA. Results from Rvachew and Grawburg (2006) confirmed that speech perception directly affects PA in children with SSD and suggested an additional mediating factor. In a large sample of preschool children aged 48-67 months, speech perception was found to have both a direct effect on PA, as well as an indirect effect through a measure of receptive vocabulary. In a second

study, conducted by Benway et al. (2021), speech perception in school-age children with residual speech /r/ errors was measured by identification of the initial sound of a *rake-wake* continuum. This measure of perceptual acuity was also directly related to PA (Benway et al., 2021).

Speech production

In standard oral PA tasks, adequate speech production skills are required to correctly imitate the presented stimuli and produce its manipulated form. It is also hypothesized that speech production is linked to phonological memory through automatic brain activation patterns. When an individual accesses knowledge of a particular phoneme, a cognitive rehearsal of the motor articulation patterns for that phoneme is also simultaneously activated (Gathercole & Adams, 1993). This suggests that accurate production of an individual phoneme supports the strength of its phonological representation and may lead to variance in a child's PA for that phoneme. It is reasonable that errored articulation patterns, then, may lead to weak perception and phonological representation, as the child is unable to correctly recognize and store the proper production (Sutherland & Gillon, 2007, Sénéchal et al., 2004).

While a direct association between articulation and speech perception is well established (Sénéchal et al., 2004; Benway et al., 2021; Sutherland & Gillon, 2007; McAllister Byun & Tiede, 2017), the association between articulatory accuracy and phoneme awareness exhibits some variation across typically developing children and those with SSD.

Sutherland and Gillon (2007) conducted one longitudinal study with preschool children with and without SSD. Children with SSD performed consistently worse than TD on speech perception and receptive phonological representation tasks. These results, and others, support the proposal that children with errored productions have greater difficulty storing underlying phonological representations (Elbro et al., 1998; Tambryraja et al., 2020; Farquharson et al., 2021).

Rvachew and Grawburg (2006) tested the hypothesis that underlying deficits in articulatory gestures lead directly to difficulties in speech perception and reading ability in children with SSD. Linear structural equation modeling revealed this as a weaker

statistical model. Articulation did not yield direct effects on PA skills. In fact, it related to PA only in that it was strongly predicted by speech perception. Results suggested that children with SSD, regardless of severity, can develop PA the same as their TD peers (Rvachew & Grawburg, 2006).

Meanwhile, investigations with children with age-expected articulation skills demonstrate rather consistent predictive correlations with PA, with most reaching statistical significance (Sénéchal et al., 2004). Erskine et al. (2020), as described above, also included a measure of speech production using the Goldman Fristoe Test of Articulation – Second Edition (GFTA-2; Goldman & Fristoe, 2000). Articulation accuracy was found to be a significant predictor of PA at age 5. Children with fewer speech sound errors at an early age (2;4-3;4) later scored significantly higher on formal measures of PA when entering kindergarten (Erskine et al., 2020).

Vocabulary

Vocabulary size has indisputable impacts on PA. The nature of this relationship is evident across levels of linguistic complexity for PA tasks. Compound words are one broad example of the support of word knowledge in understanding underlying word structure. For example, familiarity with the words *mail* and *man* lead to explicit awareness that *mailman* can be broken down into component parts. Similarly, knowledge of similarly structured words supports awareness of individual phonemes. One becomes increasingly aware of the distinct and manipulable sounds within *cat* when contrasted by words such as *cab, hat,* and *at.*

When describing gross vocabulary size, research metrics have varied between exclusively receptive or expressive measures, resulting in supposed discrepancies in the nature of the relationship between vocabulary and PA. While the following summary will discuss each respectively, the current study will consider expressive vocabulary alone.

The emergent view of phoneme development explains that each word added to a child's lexicon requires an increased level of segmentalized awareness. Knowledge of a word's unique sound structure allows discrimination from its growing neighborhood of similar-sounding words (Walley et al., 2003; Metsala, 1999). Restructuring occurs particularly to the sound structures represented within the child's vocabulary, as each

word added to the lexicon requires a level of segmentalization unique to the words already known. Therefore, vocabulary growth acts as the driving force for phonological awareness. This is evidenced by reduced vocabulary skills in dyslexic children, whose PA skills are markedly delayed when compared to normal-reading peers (Joanisse et al., 2000).

To demonstrate this relationship empirically, Metsala (1999) investigated receptive vocabulary and PA with typically developing preschoolers aged 4-6. The absolute size of a child's vocabulary was positively related to performance on PA tasks across a variety of linguistic contexts. Children with larger vocabularies consistently outperformed those with smaller vocabularies on PA tasks regardless of age. Similarly, the effects of word characteristics, such as lexical status or structure, on PA performance were more apparent for children with smaller vocabularies (Metsala, 1999).

Several studies stand in support of this direct relationship between measures of receptive vocabulary and PA (Sénéchal et al., 2004). Models proposed by Rvachew and Grawburg (2006) utilized a standardized assessment of receptive vocabulary. Participants were a cohort of preschool children ages 4-5 with primary SSD. Receptive vocabulary size directly predicted performance on PA tasks as well as mediated the effect of speech perception (Rvachew & Grawburg, 2006). Benway et al. (2021) corroborated these results, indicating that up to 19% of the variance of PA could be attributed to receptive vocabulary when combined with the measure of perceptual acuity described earlier (Benway et al., 2021). However, this model did not support receptive vocabulary as a mediating variable between speech perception and PA.

Contrasting evidence can also be found in the literature. Erskine et al. (2020) measured receptive vocabulary of a large sample of children ages 2-4. Performance on the Peabody Picture Vocabulary Test – Fourth Edition (PPVT-4; Dunn & Dunn, 2007) was not found to be significantly predictive of PA two years later. Instead, receptive language was highly correlated to performance on nonword repetition tasks which, in turn, was predictive of PA (Erskine et al., 2020). Interestingly, other studies conducted with preschool and elementary age children have found that global receptive vocabulary measures were predictive of word reading ability but not PA skills (Elbro et al., 1998; Garlock et al., 2001; Walley et al., 2003).

Few studies have considered expressive language independent from receptive vocabulary in relation to PA or associated literacy skills. Lonigan et al. (2000) considered the oral language skills of 96 children from early to late preschool years using standardized measures of single-word naming, descriptive language, and grammar. Expressive language exerted direct effects on later decoding skills as well as indirect effects through phonological sensitivity measures. Together, this accounted for up to 25% of the variance in decoding skills (Lonigan et al., 2000). However, when measured by the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Hill et al., 1985), expressive language exerted only indirect effects on PA through experience with reading-related activities or instruction (Foy & Mann, 2005).

Inconsistent correlations between receptive and expressive vocabulary with PA may be a result of inconsistent measures, as in absolute size or a standardized assessment, age of participants, or the nature of current versus predictive correlations. Nevertheless, evidence demonstrates a strong agreement of the role of vocabulary in spurring the lexical restructuring that underlies early PA and beyond.

Nonword repetition accuracy: Higher-order phonological knowledge

Nonword repetition (NWR) indexes a variety of skills, including vocabulary and phonological working memory, and presents another opportunity for predicting PA at an early age. The mechanisms required in NWR tasks to acoustically process, phonologically encode, store, and retrieve novel sound sequences resemble that of oral PA tasks. Opportunely, the ability to repeat nonwords emerges long before a child is expected to identify and isolate phonemes.

NWR tasks are designed to uniquely tap different component processes while minimizing the influence of other properties, such as word knowledge. According to the LRH, children with larger vocabularies will have more fine-grained segmental representations of words, particularly words that are highly familiar. Finer representations and the ability to isolate these representations support the ability to encode and repeat novel sequences in words and nonwords. In this model, PA can be considered a mediating variable between vocabulary and NWR accuracy (Metsala, 1999).

Research outside of the LRH continues to confirm a similar relationship between vocabulary size and NWR tasks in children as young as 2-3 years old (Gathercole & Adams, 1993). For example, low performance on NWR tasks has proven to be a salient indicator of language impairment for school-age children. Language-impaired children perform significantly worse on NWR tasks than peers with normal language skills (Dollaghan & Campbell, 1998).

Metsala (1999) designed a multifaceted investigation of PA and NWR skills of children ages 4-6. Sixty–one students completed standardized measures of receptive vocabulary, cognition, and word- and letter-identification tasks. PA was measured through rime blending and phoneme isolation tasks while NWR, short-term memory of CVC words, and digit span were combined into a latent NWR construct. Single regression models demonstrated significant correlations between PA and NWR as well as PA and receptive vocabulary. Performance on NWR tasks was influenced by both word-likeness and age, with greater effects of word likeness on younger children. Children also demonstrated stronger PA for familiar words than pseudowords. The author concluded that NWR and vocabulary were more accurately related through the interagent of sublexical processing (Metsala, 1999).

With initial evidence for word familiarity as a meaningful influence on NWR, Edwards et al. (2004) cross-sectional study considered the relationship between vocabulary size and the effect of high- or low-probability sublexical sequences. Standardized measures of receptive and expressive vocabulary were taken from 104 children ranging from ages 3;2-8;10. In addition, children repeated 44 nonwords, which were grouped into 22 matched pairs. Each pair contained a contrasting two-phoneme sequence, one of which occurs in many real words. For example, /ft/ occurs in *after* and *fifty* while /fk/ occurs only in some surnames, like *Kafka*. Production accuracy was systematically scored for nonwords and considered in relation to the frequency of the embedded diphone sequences. As hypothesized, there was a main effect of frequency of sublexical sequences on accuracy for all age groups. However, as vocabulary size increased, particularly expressive vocabulary, the effect of target sequence frequency decreased. This was true across low-frequency and zero-frequency sequences,

underscoring the role of vocabulary in the development of robust phonological representations (Edwards et al., 2004).

Erskine et al. (2020) later introduced a nonword repetition task that was designed to index two measures of phonological processing. Erskine et al.'s work was based on previous findings by Edwards et al. (2004), employing the same word and nonword stimuli as well as the same systematic coding of accuracy of the target diphone sequences. Methods were based on the hypothesis that sequences like /ft/ could be repeated by recruiting knowledge that exists in the lexicon. To repeat sequences like /fk/ accurately, children must have generalized that /f/ and /k/ can occur outside of the words in which they were learned. This task was administered at two time points: ages 2;4-3;4 (Time 1) and 4;3-5;6 (Time 2). At Time 2, the Comprehensive Test of Phonological Processing - 2nd Edition (CTOPP-2, Wagner et al. 2013) was administered as the dependent variable of PA.

Results were analyzed in a mixed-effects regression model, which included a child-specific intercept and a child-specific slope. Slope reflected how strongly sequence frequency affects repetition accuracy. Children who have phonological knowledge abstracted away from words are predicted to have a shallower slope than children whose phonological knowledge is dependent on the lexicon. The intercept reflects overall repetition accuracy; in other words, a more-general measure of phonological processing. Shallower slopes and higher intercepts were expected to correlate with larger vocabularies and stronger PA skills.

Analysis first revealed a significant correlation between vocabulary and NWR accuracy. Children more accurately produced nonwords containing high probability sequences. Regressions then demonstrated a significant effect of child-level random intercepts on PA, but not effects of child-level random slopes. The overall measure of NWR accuracy at Time 1 accounted for approximately 22% of the variance in PA skills for children two years later (Erskine et al., 2020). These results bolster the clinical use of NWR tasks with children ages 2-3 to reliably predict PA prior to kindergarten.

More recently, Cychosz et al. (2021) examined a different, and arguably more ecologically valid, measure of the extent to which children’s phonological knowledge has been generalized from the lexicon. Specifically, the study examined whether a lexicality

effect would show relationships with vocabulary size similar to the way that phonotactic probability effects did in Edwards et al. (2004). Eighty-six children completed real and nonword repetition tasks in two sessions as part of a larger study examining relationships between vocabulary growth and phonological development. Children were aged 3;3-4;4, English-speaking, with normal hearing levels and typically developing speech and language. Receptive and expressive vocabulary measures were included as part of a longer longitudinal study including these variables. These measures were included to test for interaction effects between NWR and vocabulary.

The lexicality effect was measured by comparing the accuracy of the repetition of word-initial CV sequences in real words and nonwords. Real and nonwords were scrupulously matched across levels of phonotactic transitional probability, frequency of the VC sequence following the initial consonant, frequency of second syllable, and duration of coarticulation, ensuring that only variance from the lexical status remained. Performance was ultimately measured by the accuracy of production of the initial CV sequence. Results demonstrated that children produce the CV sequences in real words more accurately than in nonwords. This effect was separate from any other potential effects (i.e., the frequency of the parts of the word other than the initial CV). Main effects of receptive and expressive vocabulary demonstrated that children with larger vocabularies produced more accurate productions regardless of word type. Similarly, the impact of word type decreased as children's vocabulary sizes increased. While some children with smaller vocabularies were able to accurately produce nonwords, children with the largest vocabularies produced some nonwords more accurately than their real word pair. Overall, results demonstrated a clear lexicality advantage in children's production accuracy for real over nonword repetition (Cychosz et al., 2021).

Previous modeling of component variables

Research continues to divulge and expand concurrent relationships between skills developing pre-PA. Early theories of PA proposed three underlying components: speech perception, STM, and general cognition. McBride-Chang's (1995) best-fit model of PA demonstrated moderately significant correlations between these variables and a PA construct. When considered together, speech perception, STM, and cognition accounted

for nearly 60% of the variance in PA. While each of these contributed uniquely to PA, only general cognitive skill and verbal memory were significantly correlated (McBride-Chang, 1995). With the processes underlying short term memory and speech perception not well understood and a known correlation between short-term memory and nonword repetition, Metsala (1999) then turned attention toward the impact of vocabulary and nonword repetition. The emergent view of spoken word recognition posits that the segmental representations of phonemes needed for basic PA emerge from global and local growth of the lexicon. For children ages 4-6, receptive and expressive measures of vocabulary were positively related to blending and phoneme isolation tasks. Results also demonstrated a positive progression of these skills with age and familiarity; Older children consistently outperformed their younger counterparts and increased performance for words acquired at a younger age (Metsala, 1999). Later studies reiterated the strength of vocabulary within the predictive framework, indicating that children with larger vocabularies also demonstrate greater accuracy in nonword repetition tasks (Edwards et al., 2004). Together, these results suggest both a direct and potential indirect effect of vocabulary on PA through nonword repetition.

Benway et al. (2021) proposed a similar relationship between vocabulary and speech perception. For school-age children with residual speech sound errors, receptive vocabulary measures also wielded direct effects on PA. Speech perception, when measured by perceptual acuity on a rake-wake continuum, was significantly correlated with PA. However, when measured as perceptual bias, the relative broadness of a perceptual category, there was not a direct relationship with PA. Though raising inconsistencies between measures of speech perception, this framework accounted for 19% of the variance in PA (Benway et al., 2021).

Two recent studies examined relations between predictive variables of PA. Erskine et al. (2020) provides one longitudinal analysis of early correlates of PA in children aged 2;4-3;4. Research sought to determine the extent to which measures of speech perception, speech production, NWR, and receptive vocabulary predicted performance on standardized measures of PA two years later. Together, these 4 variables accounted for nearly 27% of the variance in performance on the CTOPP-2 at age 5. Contrary to previous results, production-based measures yielded greater results than

perception. Significant effects were reported for both speech production and accuracy of production of nonwords, deemed child-level random intercepts. Children who produced more features accurately in nonword sequences or fewer errors on standardized articulation measures performed higher on standardized measures of PA. Meanwhile, nonsignificant effects were reported for speech perception, measured through minimal pair discrimination, and receptive vocabulary (Erskine et al., 2020). Though this framework underscores children's phonological processing and articulatory-motor control as the most significant predictors, these factors have historically produced mixed results. Clinically, these results show that nonword repetition accuracy at age 2;4-3;4 is a reliable predictor of phonological awareness 2 years later.

Rvachew and Grawburg (2006) offers a second comprehensive analysis of PA correlates in young children with speech sound disorders to support the roles of speech perception and receptive vocabulary in predicting PA. Ninety-five preschool children aged 48-67 months with primary speech sound disorder were presented with various standardized and/or previously vetted measures of articulation accuracy, receptive vocabulary, and emergent literacy skills, and computer-based speech perception tasks. Participant performance was mapped onto 2 possible relational frameworks for PA and reported results of best fit using standardized regression weights.

In the first model, hereby referred to as Model A, the externally derived variable was articulatory skill. Articulation was modeled to contribute directly to speech perception and both directly and indirectly to receptive vocabulary through speech perception. Receptive vocabulary then stands as the sole predictive contributor to PA and emergent literacy knowledge. In this proposed framework, primary deficits in articulation are identified to lead to later reading difficulties. Model B instead supposes speech perception as the exogenous variable, generating direct effects on articulation, receptive vocabulary, and PA as well as an indirect effect on PA through receptive vocabulary. Fit statistics showed a significant preference for Model B over Model A, contending that speech perception is directly related to phonological encoding and therefore later reading ability. Model B supports PA performance being predicted by speech perception and receptive vocabulary skills and reaffirms that emergent literacy knowledge is overwhelmingly explained by PA. Furthermore, the influence of speech perception on PA

is mediated by the influence that speech perception has on vocabulary size (Rvachew & Grawburg, 2006).

These results have considerable implications for PA and for children with SSD in establishing that children with SSD develop PA similar to typically developing peers. It hailed the importance of both perception- and production-based measures in predicting PA. However, the nature of linear structural equation modeling and the inconsistencies gives pause to assumptions that these models are complete. Further research is needed, in both replication and extension, to determine if these results are reproducible within and across TD and non-TD preschoolers prior to age 4.

Aims of the current study

The current study seeks to advance the theoretical understanding of the concurrent relationship between speech perception, vocabulary, NWR accuracy, speech production, and cognition in children at the approximate age of school entry. Despite evidence that measures of these respective variables yield unique and significant effects on PA individually, direct effects have varied when considered concurrently.

The following analysis will present an expanded, quasi-replication of Rvachew and Grawburg (2006). First, speech perception, speech production accuracy, vocabulary, and PA will be presented in the relationships proposed in the researcher's preferred model. Additional models will expand this framework to include two additional manifest variables: a measure of higher-level phonological knowledge, the lexicality effect described by Cychosz et al. (2021) and nonverbal cognition. The lexicality effect will examine the role of higher-level phonological knowledge while the addition of general cognitive ability will address a potential confounding variable required for the cognitive manipulation that underlies PA.

Methods

This paper presents a secondary analysis of data collected from a group of children in a longitudinal study on phonological development and vocabulary growth. Some of the data in this paper have been presented elsewhere, in articles focusing on different research questions (Erskine et al., 2021; Munson et al., 2021). This methods section provides a summary of the methods used to collect these data, with a level of detail that is appropriate for understanding the secondary analysis presented herein. Interested readers are encouraged to read the articles that expound on these data at greater length for more information.

Participants

Participants in this study were a subset of the 164 children who were recruited for the larger longitudinal study of vocabulary and phonological acquisition. These children were recruited through public advertisements near universities in Madison, WI and Minneapolis, MN. While the larger study examined multiple groups of children (including children with cochlear implants and late talkers), the 164 children from whom data were taken for this study had typical language and hearing. Participants were monolingual English speakers of Mainstream American English (MAE) or African American English (AAE) dialects. Dialect was identified after recruitment in order to provide stimuli in the child's respective dialect. Normal hearing was confirmed in at least one ear, confirmed at 25 dB hearing level for 1000, 2000, and 4000 Hz. Parent report confirmed typically developing speech and language skills and denied any current special education services. Despite efforts to recruit a sample that reflects socioeconomic diversity in the United States, an analysis of maternal education level revealed that 83% of mothers had received at least a college degree.

At the first time-point (Time 1), children were 2;4-3;4 years old. These children were followed across three time points at approximately yearly intervals. This study examines the data from the cohort of children at Time 3, when they were aged 4;4-5;4 (years;months). Candidacy for the current study was based on the completeness of children's datasets, as described below.

At Time 3, participants completed a variety of assessments across two or three one-hour sessions. The tasks outlined in the current analysis represent only a subset of measures taken by the original research. Data from 72 children were excluded due to incomplete data on at least one task of interest, leaving 92 children (49F, 43 M) for this analysis.

Tasks

Cognition

The Kaufman Brief Intelligence Test – Second Edition (KBIT-2, Naugle et al., 1993) was selected as a measure of general cognition. It serves as a briefer measure than the broader, omnibus assessments like the WISC-III, which was used in McBride-Chang's (1995) PA framework. Subtest raw scores were combined to form a composite IQ raw score. The corresponding standard score was then recorded as a measure of nonverbal intelligence.

Speech perception

Children's speech perception skills were assessed using the computer-based game Speech Assessment and Interactive Learning System (SAILS; Rvachew, 1994). Stimuli included recorded tokens from both children and adults. Words beginning with frequently misarticulated consonants (/l/, /k/, /ɹ/, /s/) were presented in conjunction with two images. While both images pictured the target word, one of the images, overlayed by a general prohibition symbol, represented an incorrect production. Children distinguished between correct (e.g., rat→/ræt/) and incorrect productions (e.g., rat→/wæt/) by choosing the corresponding image. From the 70 test trials, a percentage correct represented performance.

Nonword repetition accuracy

The nonword repetition task followed the model presented by Edwards & Beckman (2008). Children were presented with picture- and audio-prompted stimuli of 23 real and 67 nonwords. Stimuli recorded in either Mainstream American English (MAE) or African American English (AAE) dialect were presented, according to the

child's respective dialect. All words were syllable-length and distinguished by their initial CV sequence. To isolate the effect of word status on performance, words were matched methodically to control for the effects of articulatory skill, stress, prosodic structure, and phonotactic probability. In other words, only the support of existing lexical representations of real words, a lexicality effect, would support repetition accuracy.

Participant responses were recorded and later analyzed systematically using PRAAT. Trained phoneticians then transcribed the target CV sequences. The ICC between raters was 0.881 ($F(374,375) = 15.9$, $p < .001$, 95% CI = 0.86, 0.9). Consonant productions were scored based on feature analysis and overall structural accuracy when compared to the target word. Each sequence could receive a maximum of seven points: 3 points for each of the two sounds (representing accuracy of place, manner, and voicing for consonants, and height, backness, and tense/lax for vowels), and 1 point for the sequence being produced without any phoneme deletions or additions. The summary scores used in this investigation were the difference in raw accuracy between the target sequence in the real word and nonword pair. In principle, the accuracy score could range from +7 (if the real word was produced perfectly and the nonword was produced entirely inaccurately) to -7 (if the nonword was produced perfectly and the real word was produced entirely inaccurately). For more detailed information regarding verbal instructions, scoring procedure, and complete list of stimuli, see Cychosz et al. (2021).

Articulation

A standardized measure of articulation was collected through the Goldman Fristoe Test of Articulation – Second Edition (GFTA-2; Goldman & Fristoe, 2000). Participant productions were recorded and analyzed by three graduate student transcribers with background in linguistics or speech pathology. Each transcriber scored productions independently unless concerns for accuracy arose. In such cases, transcribers presented the recordings to the group for consensus scoring. Raw scores, rather than standard scores, were employed in an effort to replicate the models of Rvachew and Grawburg (2006) more closely.

Vocabulary

The Expressive Vocabulary Test - Second Edition (EVT-2; Williams, 2007), a picture-based naming assessment, provided a standardized measure of expressive vocabulary. As did Rvachew and Grawburg (2006), raw score data comprised the vocabulary construct.

Phonological awareness

Finally, phonological awareness was measured by three subtests of the Comprehensive Test of Phonological Processing – Second Edition (CTOPP-2; Wagner et al., 2013): Blending, Elision, and Memory. Various tasks measured a child's ability to synthesize or extract phonological segments to form new words. Given the high correlation observed between performance on each subtest, raw scores were recorded and transformed to a single PA construct, *PA Composite*.

Results

General procedures

Report of the modeling procedure will be as follows. First, we will present descriptive data on the measures (Table 1) and as well as correlations among them (Table 2). Second, we will examine a model commensurate with Rvachew and Grawburg (2006), including the influence of speech perception, speech production accuracy, and vocabulary on PA. The third section will examine a series of models that include these same variables as well as an expanded lexicality effect. The fourth section presents the most comprehensive model of PA: speech perception, speech production, vocabulary, lexicality effect, and nonverbal cognition.

Each section begins with a series of multiple regression analyses that examine predictors of the composite measure of phonological awareness. This presents an unstructured examination of the independent measures, distinct from the subsequent path analysis in which independent variables are structured relative to one another.

Descriptive analyses

Table 1. Group performance on all measures

Construct	Measure	M	SD	Range
Speech perception	SAILS (% correct)	0.73	0.10	0.51-0.94
Speech production	GFTA-2 (raw)	56.56	11.11	27-72
Vocabulary size	EVT-2 (raw)	75.72	16.72	24-105
Nonword repetition accuracy	Lexicality effect	0.31	0.47	-0.39-2.57
Cognition	KBIT-2 (standard)	107.65	13.34	73-145
Phonological Awareness Composite	CTOPP-2 Elision, Blending, Memory (raw)	10.51	5.61	0-26

Note: *SAILS = Speech Assessment and Interactive Learning System; GFTA-2 = Goldman Fristoe Test of Articulation – Second Edition; EVT-2 = Expressive Vocabulary Test – Second Edition; KBIT-2= Kaufman Brief Intelligence Test – Second Edition; CTOPP-2 = Comprehensive Test of Phonological Processing – Second Edition*

Table 2. Bivariate Pearson product-moment correlations among variables

	1	2	3	4	5	6
1. EVT-2	-	-	-	-	-	-
2. GFTA-2	0.371	-	-	-	-	-
3. KBIT-2	0.29	-0.023	-	-	-	-
4. SAILS	0.387	0.294	-0.003	-	-	-
5. Lexicality Effect	-0.483	-0.279	-0.173	-0.204	-	-
6. PA Composite	0.681	0.477	0.227	0.349	-0.508	-

Note: Anything whose absolute value is 0.267 or more is significant at p<0.01. Anything whose absolute value is 0.205 or less is significant at p<0.05. EVT-2 = Expressive Vocabulary Test–Second Edition raw score; GFTA-2 = Goldman Fristoe Test of Articulation–Second Edition raw score; KBIT-2= Kaufman Brief Intelligence Test – Second Edition standard score; SAILS= Speech Assessment and Interactive Learning System % correct; PA Composite= Comprehensive Test of Processing Phonological composite raw scores (Blending, Elision, Memory subtests)

Model 1: Influence of speech perception, speech production, and vocabulary on PA

The first step in the analysis was to examine predictors of the PA Composite from speech perception (SAILS), speech production accuracy (GFTA-2), and expressive vocabulary (EVT-2) using multiple regression. This regression was significant, $F [3,88] = 26.9$, $p < .001$, multiple $R^2 = .48$. As seen in Table 3, the coefficients for EVT-2 and GFTA-2 were significant, while the variable for speech perception was not.

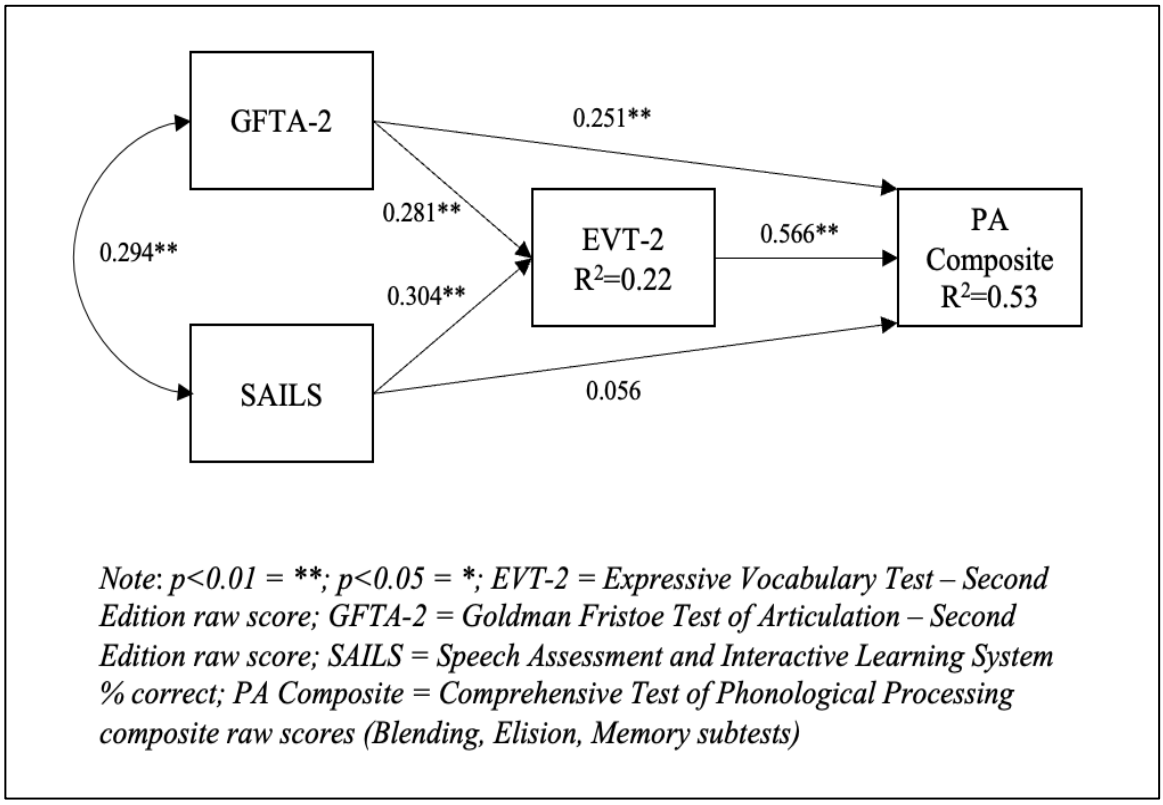
Table 3. Regression Model 1 Coefficients

	Estimate	SE	T Value	P-Value
Intercept	-22.88	7.48	-3.06	0.003
EVT-2	0.43	0.06	6.80	<0.001
SAILS	0.07	0.10	0.69	0.49
GFTA-2	0.29	0.09	3.12	0.002

Note: EVT-2 = Expressive Vocabulary Test – Second Edition raw score; GFTA-2 = Goldman Fristoe Test of Articulation – Second Edition raw score; SAILS = Speech Assessment and Interactive Learning System % correct

The next set of analyses (Figure 2) examined whether speech perception and speech production influence PA directly, or whether their influence is solely based on the mediating role of vocabulary size. To examine this, we constructed a path analysis, similar to that in Benway et al. (2021). As is true with Benway et al., this model is saturated, meaning that the overall model fit cannot be assessed. However, the individual coefficients within the model are nonetheless interpretable. As the coefficients below show schematically, there were direct effects of EVT-2 and GFTA-2 on PA. There were also effects of GFTA-2 and SAILS on EVT-2 as well as significant covariance between them. This means that, for this data set, all three variables predict PA. For GFTA-2, this effect is both direct and indirect through EVT-2. For SAILS, this effect is solely indirect, through EVT-2.

Figure 2. Path Model 1, including speech perception, speech production, and vocabulary size



Model 2a and 2b: Lexicality effect as a unique predictor of PA

The next analyses examined models (Figure 3; Figure 4) that included a third phonological measure, lexicality effect. Recall that this represents the difference in the accuracy with which word-initial CV sequences were repeated when they were in real words (i.e., /su/ in *suitcase*) or nonwords (i.e., /su/ in *sudrass*). Cychosz et al. (2020) argued that this is a higher-level phonological skill, indicating the extent to which children's phonological knowledge can be rapidly generalized to unfamiliar words.

As a first step in examining the potential role of the lexical effect, we ran a multiple regression (Table 4) including it as a predictor, along with speech perception, speech production, and vocabulary size. This regression was significant, $F\ [4,87] = 23.9$, $p < .001$, multiple $R^2 = .523$.

Table 4. Regression Model 2 Coefficients

	Estimate	SE	T Value	P-Value
Intercept	-14.41	8.04	-1.79	0.08
EVT-2	0.37	0.07	5.37	<0.001
SAILS	0.07	0.10	0.71	0.48
GFTA-2	0.26	0.09	2.89	0.005
Lexicality Effect	-0.06	0.02	-2.47	0.02

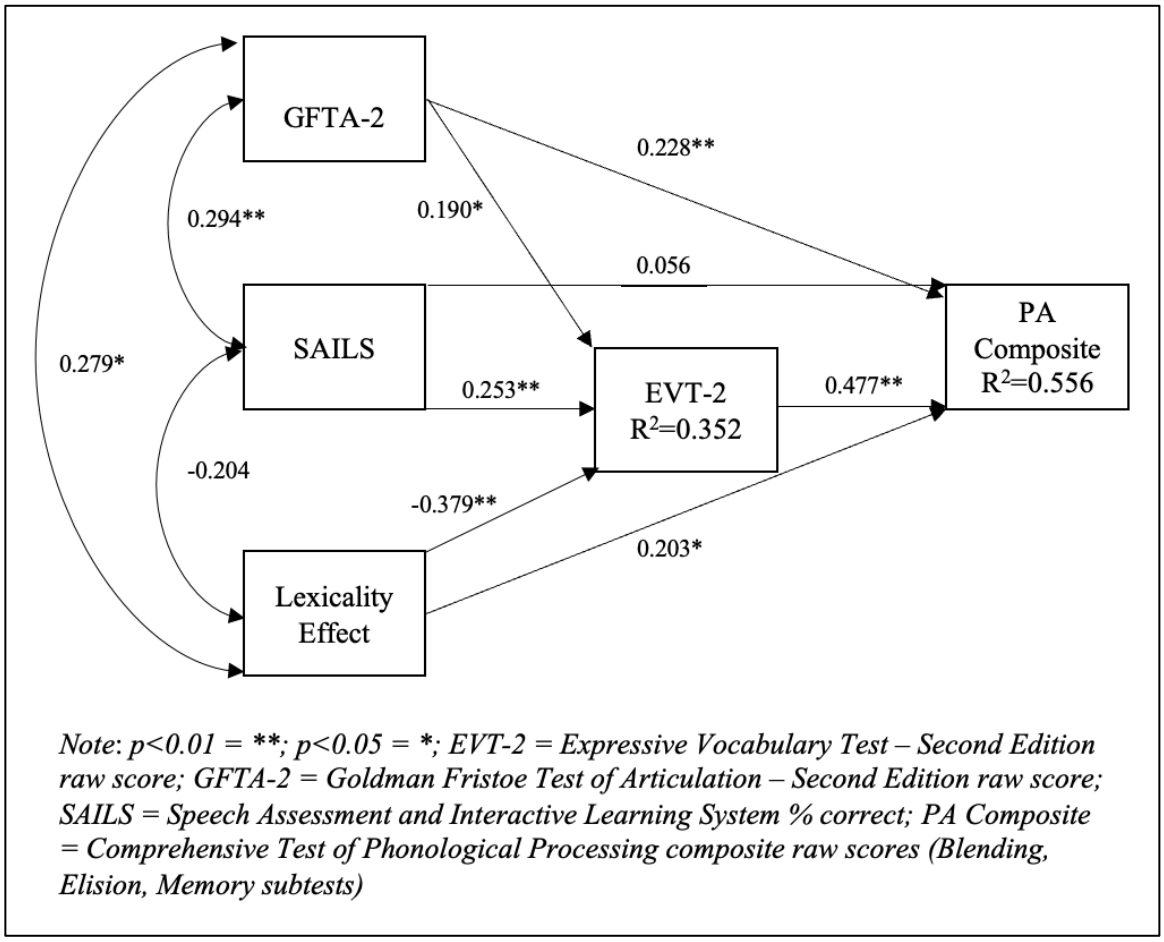
Note: EVT-2 = Expressive Vocabulary Test – Second Edition raw score; GFTA-2 = Goldman Fristoe Test of Articulation – Second Edition raw score; SAILS = Speech Assessment and Interactive Learning System % correct

Two path analyses were then conducted that included lexicality. The first of these, Figure 3, treated lexicality effect simply as another measure of phonological ability which could predict PA both directly, and indirectly through EVT-2. That model was again saturated, meaning that its interpretation depends not on overall model fit, but on the significance of different components of the model.

As this figure shows, the addition of lexicality effect in Path Model 2a did not change the overall patterns of the influence of GFTA-2, SAILS, and EVT-2 on PA: both

speech production and speech perception continued to predict vocabulary size, which in turn predicted PA. Moreover, there continued to be direct effects of GFTA-2 on PA, but not SAILS. Lexicality effect had both a significant effect on EVT-2, and a direct effect on PA. It also covaried significantly with GFTA-2, but not with SAILS. The addition of lexicality effect in the model increased the variance in EVT-2 that was accounted for considerably. The variance accounted for in PA also increased, albeit more modestly.

Figure 3. Path Model 2a, including speech perception, speech production, vocabulary size, and lexicality effect

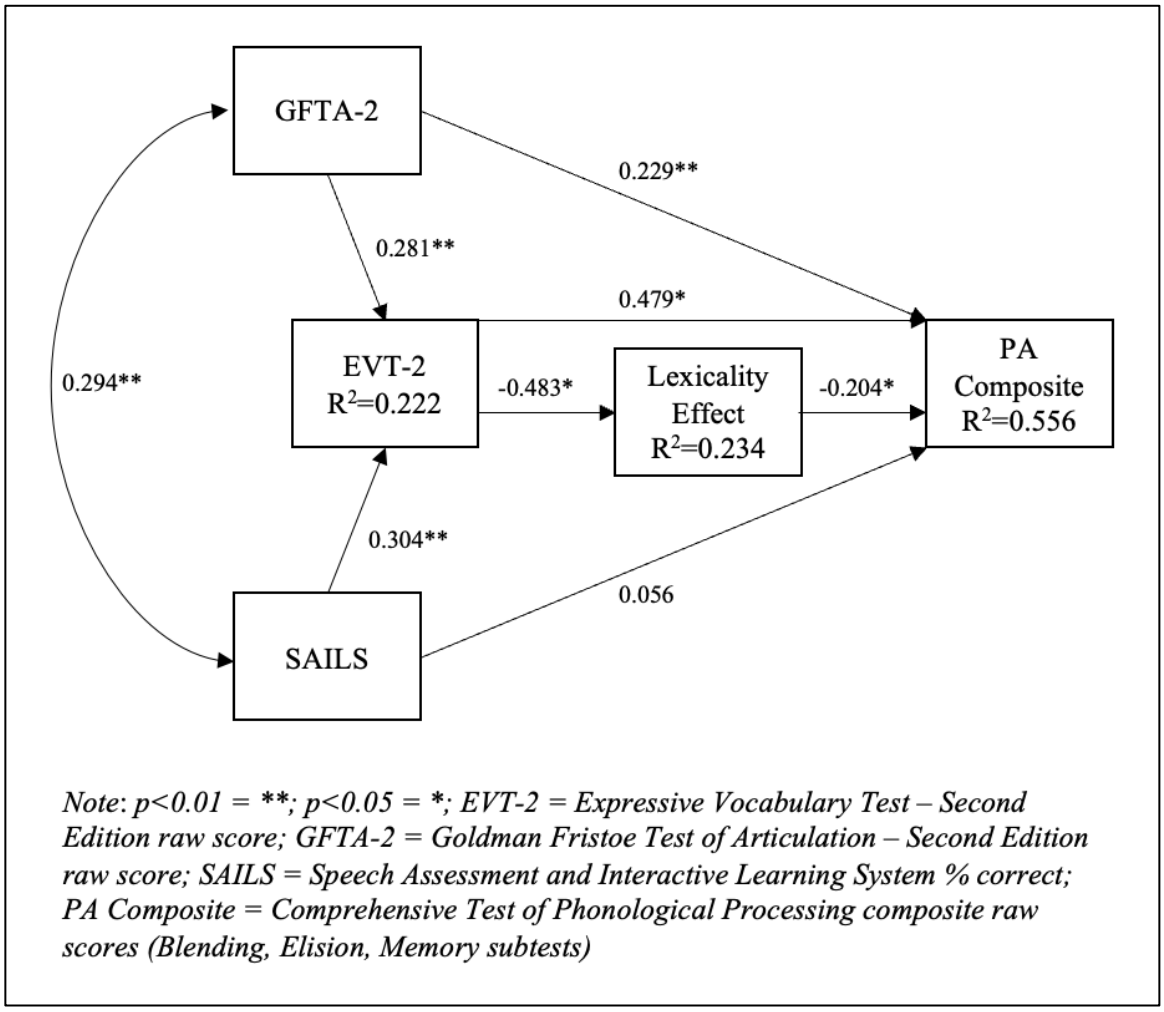


Note: *p<0.01 = **; p<0.05 = *; EVT-2 = Expressive Vocabulary Test – Second Edition raw score; GFTA-2 = Goldman Fristoe Test of Articulation – Second Edition raw score; SAILS = Speech Assessment and Interactive Learning System % correct; PA Composite = Comprehensive Test of Phonological Processing composite raw scores (Blending, Elision, Memory subtests)*

The next model (Figure 4) presented the same variables in an original path structure. This different structure is motivated by theories of the development of phonological knowledge in different domains. Edwards et al. (2004) posited that higher-level phonological knowledge, of the sort indexed by lexicality effect, emerges from generalizations unique to the growing lexicon. As previously modeled, early lexical development is theorized to be supported by speech perception and speech production. Therefore, a model that reflects that structure has direct paths from speech perception and speech production to vocabulary size, from vocabulary size to the lexicality effect, and from the lexicality effect to PA. As seen below, such a model would also have direct paths from speech perception, speech production, and vocabulary size to PA.

This model was saturated. As in Path Model 2a, no changes were observed in overall patterns of the influence of GFTA-2, SAILS, and EVT-2 on PA. Expressive vocabulary, EVT-2, had both a significant influence on lexicality effect, and a direct effect on PA. Meanwhile, EVT-2, as the sole predictor, contributed a notable variance in lexicality effect. Finally, lexicality effect continued to have a significant direct effect on PA. That is, as the effect of lexical status decreased, performance on PA tasks increased.

Figure 4. Path Model 2b, including speech perception, speech production, vocabulary size, and lexicality effect



Model 3: Nonverbal cognition as a unique predictor of PA

The final model examined whether nonverbal intelligence, measured by the KBIT-2, further predicted PA. Prior to conducting path analysis, a multiple regression was fit in which the KBIT-2 standard score was an additional predictor. KBIT-2 was not a significant predictor in this model (Table 5), whose R^2, 0.529, was very similar to the model without KBIT-2.

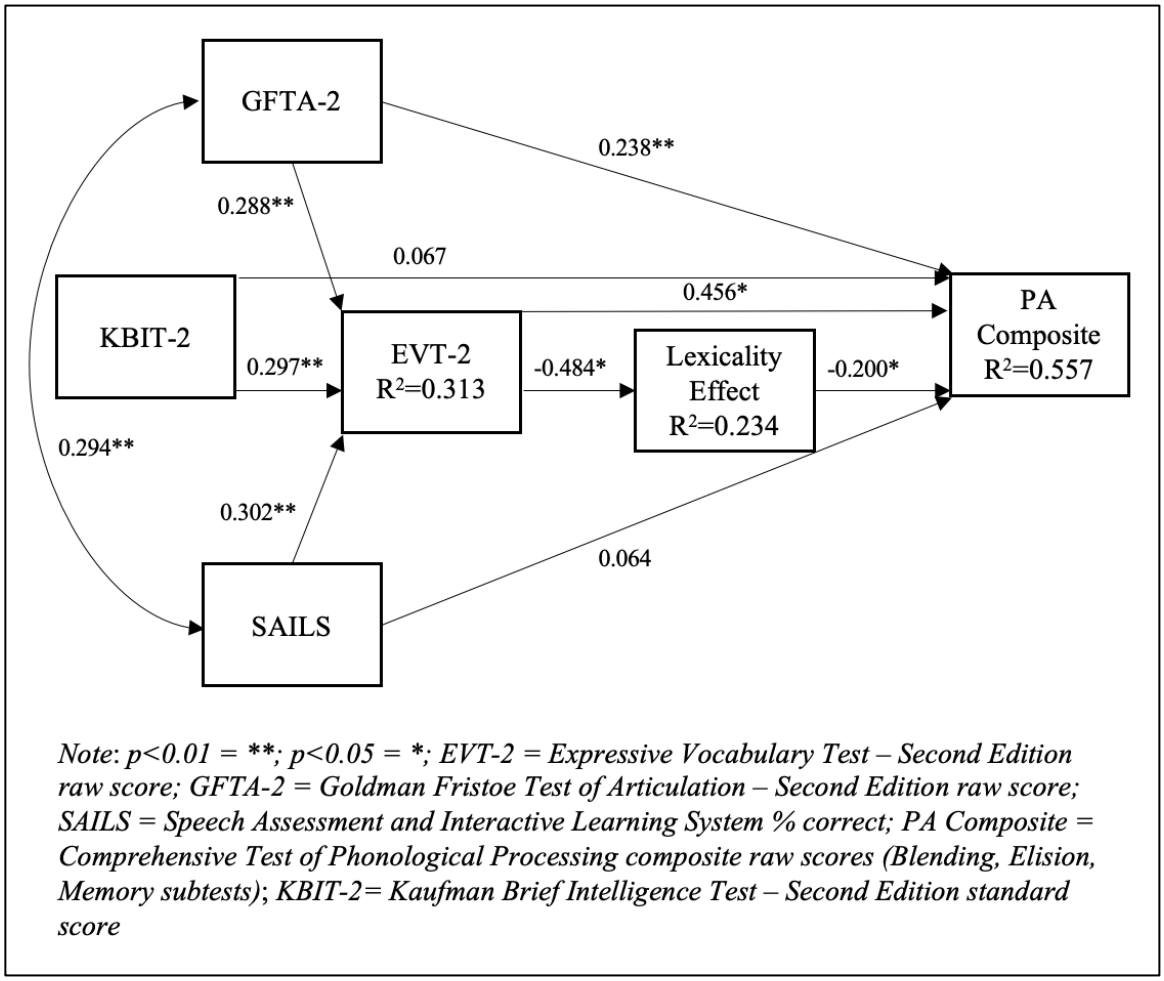
Table 5. Regression Model 3 Coefficients

	Estimate	SE	T Value	P-Value
Intercept	-20.67	10.79	-1.92	0.06
EVT-2	0.35	0.07	4.88	<0.001
SAILS	0.08	0.10	0.80	0.43
GFTA-2	0.27	0.09	2.98	0.003
Lexicality Effect	-0.05	0.02	-2.41	0.02
KBIT-2	0.06	0.07	0.87	0.39

Note: EVT-2 = Expressive Vocabulary Test – Second Edition raw score; GFTA-2 = Goldman Fristoe Test of Articulation – Second Edition raw score; SAILS = Speech Assessment and Interactive Learning System % correct; KBIT-2 = Kaufman Brief Intelligence Test – Second Edition standard score

A path analysis (Figure 5) was conducted in which the KBIT-2 score served as an independent predictor of PA. Given the well-documented relationship between measures of nonverbal cognition and vocabulary size, that model also included a path between KBIT-2 and EVT-2. As seen below, the path between KBIT-2 and EVT-2 was significant, replicating previous findings that nonverbal cognition affects vocabulary size. The influence of KBIT-2 on PA was not significant. Nor did the addition of KBIT-2 to the model change the patterns of significance for the influence of GFTA-2 and lexicality on PA.

Figure 5. Path Model 3, including speech perception, speech production, vocabulary size, lexicality effect, and nonverbal cognition



Note: *p<0.01 = **; p<0.05 = *; EVT-2 = Expressive Vocabulary Test – Second Edition raw score; GFTA-2 = Goldman Fristoe Test of Articulation – Second Edition raw score; SAILS = Speech Assessment and Interactive Learning System % correct; PA Composite = Comprehensive Test of Phonological Processing composite raw scores (Blending, Elision, Memory subtests)*; *KBIT-2= Kaufman Brief Intelligence Test – Second Edition standard score*

Discussion

In the current study, we sought to examine several theoretically supported components of PA, specifically speech perception, vocabulary, lexicality effect, speech production, and cognition. This post-hoc analysis was completed with data from a large cohort of children aged 4;4-5;4 within a longitudinal study on phonological development and vocabulary growth. On the foundation of Rvachew and Grawburg (2006), the current study presents four concurrent models of PA with several noteworthy outcomes. First, results suggest that speech perception may not be the strongest indicator of PA skills. Speech perception contributed to variance in PA only in indirect effects through vocabulary and its consistent correlation with speech production. Meanwhile, speech production accuracy robustly predicted PA in all the models that were examined. Raw score performance on a standardized articulation measure yielded significant direct effects on PA as well as indirect effects through vocabulary. Next, the lexicality effect both strengthened and expanded existing models of PA. Lexicality contributed appreciable variance, accounting for an additional 3-4%, to the PA model. Finally, for these typically developing children, nonverbal cognition does not predict PA directly, and predicts it indirectly only through its influence on vocabulary size. For contextual grounding, the following sections will expand on these findings within each respective model.

Path Model 1

Path Model 1, adapted from Rvachew and Grawburg (2006), aimed to reconsider the relationships between speech perception, speech production, vocabulary, and PA. Corroborating the original results, these three variables accounted for 53% of total variance in PA. The current study explains a modest 2% more of that variance in our PA composite. Perhaps expressive vocabulary is a stronger predictor of PA skills in five-year-old children. Or perhaps expressive vocabulary is more closely linked to the PA tasks employed in the current study. Tasks on the CTOPP-2 are largely production-based (e.g., what word do these sounds make: *cow, boy*) while Rvachew and Grawburg (2006) employed a receptive matching task.

Vocabulary size, measured by a standardized expressive vocabulary test, strongly predicted performance on PA tasks. These results substantiate previous research supporting expressive vocabulary (Metsala, 1999; Joanisse et al., 2000; Lonigan et al., 2000; Foy & Mann, 2005) as a driving force for the lexical restructuring required for the development of PA. This, in combination with significant arguments on behalf of receptive vocabulary (Elbro et al., 1998; Garlock et al., 2001; Walley et al., 2003; Sénéchal et al., 2004; Rvachew & Grawburg, 2006), validates the emergent view of phoneme development. Segmentalized awareness of the underlying structure of words grows with each unique word added to the lexicon.

In relation to the vocabulary construct, speech perception and speech production accounted for up to 22% of the variance in expressive language skill. Again, vocabulary mediated the relationship between sensitivity to the underlying structure of words as well as speech sound accuracy on PA (Rvachew & Grawburg, 2006). Of note, a strong covariance between speech production and speech perception was also maintained throughout the models. Children with more accurate auditory discrimination of phoneme-level errors also demonstrated fewer speech sound errors. This relationship is supported throughout the literature (Sénéchal et al., 2004; Benway et al., 2021; Sutherland & Gillon, 2007; McAllister Byun & Tiede, 2017).

Though both speech production and speech perception generated direct effects on vocabulary, they did not demonstrate the same effects on PA. First, our expanded model considered a direct relationship between speech production and PA. While Rvachew and Grawburg (2006) did not examine this relationship, several studies contend that speech production is related to PA, reading risk, and other variables of interest within the current model (Tambryraja et al., 2020; Farquharson et al., 2021). The current results confirm that weak phonological representations may be at least partially due to difficulty producing those particular sounds. Results from Erskine et al. (2020) suggest that this relationship is predictive of PA skills even in younger ages, when speech sound accuracy is still developing.

Surprisingly, speech perception did not yield a significant direct effect on PA. These results stand in contention with Rvachew and Grawburg (2006), among others (McBride-Chang, 1995; Benway et al., 2021), who propose speech perception as a

principal component of PA. This relationship is theoretically sound, in that the ability to accurately perceive and discriminate a sound is a prerequisite for developing a representation for that sound within the phonological memory.

One potential cause for differences lies in measures employed. Benway et al. (2021) employed a word level categorical perception task, asking children to distinguish between *rake* and *wake*. McBride-Chang (1995) presented three computer-based forced-choice perception tasks with multiple manipulated variables, including speech rate, voice onset time, and manner of articulation. Agreeable to Erskine et al. (2020), who found a lack of predictive power of speech perception on PA when using a minimal pair discrimination task, in that it may be that certain types of speech perception tasks tap into unique aspects of phonological knowledge. Historically, measures of speech perception present variably when compared to various speech-related skills, suggesting that although they reflect a single construct, tasks varying in difficulty and familiarity impact performance accordingly (McBride-Chang, 1995).

However, this reason does not explain away contention with Rvachew and Grawburg (2006), given that both studies examined word identification via SAILS. A second potential reason for variation is in the population examined. Rvachew and Grawburg (2006) collected data from a large cohort of children (age 4;0-5;7), identified with primary SSD (also more representative of SES/maternal education), while the current study recruited a broad sample of children (age 4;4-5;4) whose vocabulary, speech production, and speech perception fell within the normative range. Children with typically developing articulation skills may rely less heavily on perception-based knowledge than their counterparts with SSD.

Path Models 2a & 2b

To examine the correlation presented by Cychosz et al. (2020), Path Models 2a and 2b contrasted views of a lexicality effect as predicted *by* or predictive *of* expressive vocabulary. Path Model 2a presented the lexicality effect as an exogenous variable, predictive of performance on both PA and vocabulary measures. In this model, the lexicality effect bears strong correlations with speech production but not speech perception; though, it did not influence either correlation in respect to PA. When

accounting for the effects of speech production and speech perception, the lexicality effect contributed an additional 13% of the variance in expressive vocabulary. As vocabulary size increased, the effect of nonword status on NWR accuracy decreased, confirming the lexicality effect as a unique measure of context-independent, phonological knowledge.

Path Model 2b presented a converse relationship with similar but stronger effects. With expressive vocabulary alone predicting the lexicality effect, up to 24% of the variance in the lexicality construct was accounted for. Expressive vocabulary maintained a significant direct relationship with PA. Meanwhile, the extent to which lexical status impacted NWR performance resulted in direct variation in the PA construct. That is, children who were more successful in transferring familiar phonological representations to unfamiliar contexts, a novel word, were more likely to perform higher on PA tasks. This process is ushered directly by increases in vocabulary size, as each growing phonological representation is generalized and reinforced within the context of a newly mastered word (Edwards et al., 2004).

Notwithstanding, the addition of the lexicality effect strengthened the variance of PA within the model. Despite only modest contribution to PA, the procedure outlined in isolating lexical status as a predictive factor of NWR accuracy and later PA is a particularly unique contribution to the literature. Replication of this controlled measure with children of varying ages and demographic characteristics is sure to add meaningful contributions to the PA framework across development. Moreover, this result endorses a continued investigation into predictors of PA that represent discrete levels of phonological knowledge and develop according to individual growth in the lexicon (Edwards et al., 2011).

Path Model 3

Path Model 3 examined the role of nonverbal cognition, as measured by the KBIT-2 (Naugle et al., 1993). When placed in a developmental model, nonverbal cognition demonstrated significant direct effects on vocabulary. Children with lower language skills, on average, also demonstrated lower cognitive skills than their typically developing counterparts (Dollaghan & Campbell, 1998). Cognition did not, however,

yield direct effects on PA nor contribute significant variance to the PA construct. This was a surprising result, given the concurrent validity of the KBIT-2 and WISC-III, the measure employed by McBride-Chang (1995). Instead, the current model strengthens the argument for speech production and speech perception as unique, contributing factors of PA for children with normal nonverbal intelligence. Similar evidence presenting no to minimal correlations with related variables (Gathercole & Adams, 1993) or dyslexia (Elbro et al., 1998) presents cognition as a dormant agent in the development of phonological representations.

Summary and Conclusions

It is important to note both variation and limitation in the modeling type employed in the current study. Rvachew and Grawburg (2006) utilized structural equation modeling to view interactions between the latent and manifest measures of interest. For the current study, the latent variable, PA Composite, was transformed into a single construct. This simplified the level of regression needed and allowed for the use of path analysis as an efficient and practicable examination of the current models.

However, an inherent weakness of path analysis is the inability to determine best fit. While our models were saturated, that is not to say that one model was a better fit for the data than another nor that there is no other schema that may better represent the data. While interpretation of the regression coefficients is sound, we also cannot make claims as to the practicability of the predicted relationships proposed herein.

While our variables accounted for up to 56% of the variance in PA in five-year-old children, there are also variables of interest not considered in the models. Of note, rapid automatized naming (RAN) is one potential predictor of reading success not discussed in the current study. Research suggests that RAN represents an independent source of reading difficulty from phonological deficits (Wolf & Bowers, 2000). RAN tasks measure the speed and accuracy with which children name sequences of familiar ‘things’ (i.e., objects, numbers, colors). Logically, a task in which a child is required to promptly and effortlessly access familiar information resembles that which is required when decoding the phonological components of unfamiliar words. RAN has been proposed as a correlative measure with both vocabulary and articulation, in an

automaticity-sense. To determine, then, if RAN predicts unique variation on PA apart from speech production, future research should consider the influence of speech production accuracy and RAN on PA skills.

Limitations of the current study call for replication of the current, or expanded, models with populations of children similar or more representative than ours (i.e., children with SSD). Similarly, longitudinal predictive models of the current variables may present additional meaningful knowledge of the direct and mediating relationships of PA from an earlier age. For example, perhaps measures of speech perception and speech production may have variable effects on PA at crucial stages of language and reading development (e.g., prior to first words, vocabulary explosion around 18 months of age, explicit knowledge of morphological representations, etc.).

Yet, the proposed models confirm the operative roles of speech production, vocabulary, and lexicality effect as leading predictors of PA in young children. Meanwhile, nonverbal cognition and speech perception were found to have limited influence on PA skills. The current analysis spurs speech pathologists to target speech production and vocabulary skills at an early age. Children with speech perception and cognition across the wide normative range appear to have equal likelihood of developing robust PA. Rather, the ability to correctly produce speech sounds within an expanding neighborhood of words directly supports the development of phonological representations. For children with identified or expected reading difficulties, these early-emerging skills stand as catalysts for later reading success.

References

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, *14*(5), 255–259. https://doi-org.ezp2.lib.umn.edu/10.1111/j.0963-7214.2005.00376.x

Benway, N. R., Garcia, K., Hitchcock, E., McAllister, T., Leece, M. C., Wang, Q., & Preston, J. L. (2021). Associations between speech perception, vocabulary, and phonological awareness skill in school-aged children with speech sound disorders. *Journal of Speech, Language & Hearing Research*, *64*(2), 452–463. https://doi-org.ezp1.lib.umn.edu/10.1044/2020_JSLHR-20-00356

Brown, L., Sherbenou, R., & Johnsen, S. (1990). *Test of Nonverbal Intelligence 2*. Austin, TX: Pro-Ed.

Cychosz, M., Erskine, M., Munson, B., & Edwards, J. (2021). A lexical advantage in four-year-old children's word repetition. *Journal of Child Language, 48*(1), 31-54. https://doi.org/10.1017/S0305000920000094

Dollaghan C., & Campbell T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language & Hearing Research*, *41*(5), 1136–1146. https://doi-org.ezp1.lib.umn.edu/10.1044/jslhr.4105.1136

Dunn, M. L., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Minneapolis, MN: Pearson Assessments.

Edwards, J., & Beckman, M. E. (2008). Methodological questions in studying consonant acquisition. *Clinical Linguistics & Phonetics, 22*(12), 937–956. https://doi.org/10.1080/02699200802330223

Edwards J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language & Hearing Research*, *47*(2), 421–436. https://doi-org.ezp1.lib.umn.edu/10.1044/1092-4388(2004/034)

Edwards, J., Munson, B., & Beckman, M. (2011). Lexicon–phonology relationships and dynamics of early language development – a commentary on Stoel-Gammon's 'Relationships between lexical and phonological development in young children'. *Journal of Child Language, 38*(1), 35-40. https://doi.org/10.1017/S0305000910000450

Elbro, C., Borstrøom, I., & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: the importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly*, *33*(1), 36–60. https://doi-org.ezp1.lib.umn.edu/10.1598/RRQ.33.1.3

Erskine, M. E., Munson, B., & Edwards, J. R. (2020). Relationship between early phonological processing and later phonological awareness: Evidence from nonword repetition. *Applied Psycholinguistics*, *41*(2), 319–346. Cambridge Core. https://doi.org/10.1017/S0142716419000547

Farquharson, K., Hogan, T., & Fox, A. (2021). Factors that influence non-word repetition performance in children with and without persistent speech sound disorders. *International Journal of Language & Communication Disorders, 56*(6), 1218-1234. https://doi.org/10.1111/1460-6984.12663

Foy, J. G., & Mann, V. (2003). Home literacy environment and phonological awareness in preschool children: Differential effects for rhyme and phoneme awareness. *Applied Psycholinguistics*, *24*(1), 59–88. Cambridge Core. https://doi.org/10.1017/S0142716403000043

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language*, *45*(3), 468–492. https://doi.org/10.1006/jmla.2000.2784

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

Goldman, R., & Fristoe, M.(2000). *Goldman-Fristoe Test of Articulation* (2nd ed.). Circle Pines, MN: American Guidance Service.

Hill, T. D., Reddon, J. R., & Jackson, D. N. (1985). The factor structure of the Wechsler scales: A brief review. *Clinical Psychology Review, 5*(4), 287–306. https://doi.org/10.1016/0272-7358(85)90009-1.

Joanisse, M. F., Manis, F. R., Keating, P., & Seidenberg, M. S. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of Experimental Child Psychology*, *77*(1), 30–60. https://doi.org/10.1006/jecp.1999.2553

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*, *36*(5), 596–613. https://doi.org/10.1037/0012-1649.36.5.596

McAllister Byun, T., & Tiede, M. (2017). Perception-production relations in later development of American English rhotics. *PLoS ONE*, *12*(2), 1–22. https://doi-org.ezp1.lib.umn.edu/10.1371/journal.pone.0172022

McCarthy, D. (1970). McCarthy scales of children's abilities. New York: Psychological Corporation.

McBride-Chang, C. (1995). What is phonological awareness? *Journal of Educational Psychology*, *87*(2), 179–192. https://doi.org/10.1037/0022-0663.87.2.179

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *91*(1), 3–19. https://doi.org/10.1037/0022-0663.91.1.3

Miller, G. A. & Nicely, P. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27*, 338–352. doi:10.1121/1.1907526

Munson, B., Logerquist, M. K., Hyuna Kim, Martell, A., & Edwards, J. (2021). Does early phonetic differentiation predict later phonetic development? Evidence from a longitudinal study of /ɹ/ development in preschool children. J*ournal of Speech, Language, & Hearing Research, 64*(7), 2417-2437. https://doi-org.ezp1.lib.umn.edu/10.1044/2021_JSLHR-20-00555

National Institute of Child Health and Human Development (U.S.). (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* National Institute of Child Health and Human Development, National Institutes of Health.

Naugle, R. I., Chelune, G. J., & Tucker, G. D. (1993). Validity of the Kaufman Brief Intelligence Test. Psychological Assessment, 5(2), 182–186. https://doi.org/10.1037/1040-3590.5.2.182

Petscher, Y., Cabell, S., Catts, H., Compton, D., Foorman, B., Hart, S., Lonigan, C., Phillips, B., Schatschneider, C., Steacy, L., Terry, N., & Wagner, R. (2020). How the science of reading informs 21st-century education. *Reading Research Quarterly, 0*(0),1-16. https://doi.org/10.1002/rrq.352.

Roberts, J., Jurgens, J., & Burchinal, M. (2005). The role of home literacy practices in preschool children's language and emergent literacy skills. *Journal of Speech, Language & Hearing Research*, *48*(2), 345–359. https://doi-org.ezp2.lib.umn.edu/10.1044/1092-4388(2005/024)

Rvachew, S. (1994). Speech perception training can facilitate sound production learning. *Journal of Speech and Hearing Research, 37*(2), 347–357. https://doi.org/10.1044/jshr.3702.347

Rvachew, S., & Grawburg, M. (2006). Correlates of phonological awareness in preschoolers with speech sound disorders. *Journal of Speech, Language & Hearing Research*, *49*(1), 74–87. https://doi-org.ezp1.lib.umn.edu/10.1044/1092-4388(2006/006)

Sénéchal, M., Ouellette, G., & Young, L. (2004). Testing the concurrent and predictive relations among articulation accuracy, speech perception, and phoneme awareness. *Journal of Experimental Child Psychology*, *89*(3), 242–269. https://doi.org/10.1016/j.jecp.2004.07.005\

Sutherland, D., & Gillon, G. T. (2007). Development of phonological representations and phonological awareness in children with speech impairment. *International Journal of Language & Communication Disorders*, *42*(2), 229–250. https://doi-org.ezp1.lib.umn.edu/10.1080/13682820600806672

Tambyraja, S., Farquharson, K., & Justice, L. (2020). Reading risk in children with speech sound disorder: Prevalence, persistence, and predictors. *Journal of Speech, Language, and Hearing Research,63*(11), 3714-3726.

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*(4), 1067–1084. https://doi-org.ezp2.lib.umn.edu/10.1111/j.1467-8624.2004.00726.x

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Pearson, N. A. (2013). Comprehensive Test of Phonological Processing–2nd ed. (CTOPP-2). Austin, TX: Pro-Ed.

Wagner, R. K., Zirps, F. A., Edwards, A. A., Wood, S. G., Joyner, R. E., Becker, B. J., Liu, G., & Beal, B. (2020). The Prevalence of Dyslexia: A New Approach to Its Estimation. *Journal of Learning Disabilities*, *53*(5), 354–365. https://doi.org/10.1177/0022219420920377

Walley, A. C., & Flege, J. E. (1999). Effect of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics*, *27*(3), 307–332. https://doi.org/10.1006/jpho.1999.0098

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading & Writing*, *16*(1/2), 5–20. https://doi-org.ezp1.lib.umn.edu/10.1023/A:1021789804977

Williams, K. (2007). *Expressive Vocabulary Test –Second Edition*. Pearson.

Wolf, M., & Bowers, P. (2000). Naming-speed processes and developmental reading disabilities: An introduction to the special issue on the double-deficit hypothesis. *Journal of Learning Disabilities,33*(4), 322-324. https://doi.org/10.1177/002221940003300404