Novel Word Learning in Second Grade Predicts Later Vocabulary and Reading Comprehension

Supraja Srikumar, Anna Ehrhorn, Taylor Bryant, Dawna Duff, and Suzanne Adlof

1. Introduction

Reading comprehension is a central goal of literacy development, and vocabulary knowledge is widely recognized as one of its strongest predictors (Language and Reading Research Consortium & Logan, 2017; Perfetti & Stafura, 2014). Children who know more words tend to comprehend text more successfully, and vocabulary instruction and assessments are therefore a major focus of educational practice. Despite the strong association between vocabulary knowledge and comprehension, the extent to which individual differences in vocabulary learning ability contribute to generalized reading comprehension remains unclear.

Vocabulary knowledge itself is typically assessed using static measures that provide a snapshot of accumulated lexical knowledge at a single point in time. Such tests sample a range of vocabulary items that are calibrated for difficulty and, when well designed, provide reliable and valid estimates of a child's general vocabulary level. Children who score higher on these measures are likely to know more words overall and, on average, tend to perform better on measures of text comprehension. At the same time, standardized vocabulary measures have been criticized for indexing primarily the breadth of word knowledge and failing to capture its depth or lexical complexity (Duff et al., 2019; Hadley & Dickinson, 2020). In addition, these static scores are influenced by complex factors such as underlying learning ability, prior language exposure, instructional experiences, and socioeconomic context (Peña et al., 2016; van der Kleij et al., 2022). Because these factors shape both vocabulary growth and reading development, they may

* Supraja Srikumar, Taylor Berrier, and Suzanne M. Adlof, Department of Communication Sciences and Disorders, University of South Carolina, Columbia, SC; Anna Ehrhorn, Speech, Language, and Hearing Sciences, Auburn University, Auburn, AL; Dawna Duff, Speech and Language Pathology, Binghamton University, Binghamton, NY. Research reported in this manuscript was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under Award NIDCD R01DC017156 (PI: Adlof). We gratefully acknowledge the participating children, families, teachers, and schools, and thank Dr. Lisa Fitton for statistical consultation for this work. Corresponding author: SADLOF@mailbox.sc.edu.

 Proceedings of the 50th annual Boston University Conference on Language Development, ed. Romi Hill et al., 764-774. Somerville, MA: Cascadilla Press.

obscure individual differences in how efficiently children acquire new lexical knowledge.

Importantly, static measures do not provide insight into the process of lexical acquisition. They do not indicate how long it took a child to acquire their existing vocabulary, what learning opportunities they have had, or how efficiently they can learn new words when exposure is equated. Two children with similar vocabulary scores may differ substantially in their capacity to learn from instruction and from text. A child with rich exposure may perform well despite modest learning efficiency, whereas another child with limited exposure may perform more poorly despite a strong capacity to learn new words when given appropriate input. Distinguishing between accumulated knowledge and learning capacity requires assessments that measure acquisition under controlled conditions. Dynamic approaches to assessment aim to approximate learning processes more directly by measuring how efficiently children acquire new information when instructional exposure is equated (Camilleri & Botting, 2013; Pena et al., 2001; Petersen et al, 2020).

One such approach uses novel word learning paradigms, in which children are taught new words under controlled instructional conditions and then tested on their ability to recall and recognize them. Such tasks are designed as experimental proxies for learning efficiency. When exposure is equated across individuals, differences in performance may reflect differences in learning processes rather than differences in prior experience. Some prior work suggest that experimental measures of word learning are related to later language outcomes. Gellert and Elbro (2013) demonstrated that early word learning performance predicts subsequent vocabulary growth, and more recent work indicates that dynamic word learning tasks are associated with later comprehension development (Nash et al., 2025). Still, few studies have examined whether novel word learning ability predicts vocabulary growth longitudinally.

The present study addresses this gap by examining whether performance on a second-grade novel word learning task predicts fourth-grade vocabulary and reading comprehension above and beyond baseline receptive vocabulary, nonverbal intelligence, and socioeconomic status (SES). Using a high-context word learning paradigm and a composite index of word learning performance, we evaluate whether individual differences in word learning provide unique predictive information about later language and literacy outcomes.

2. Research Questions

The present study examined the extent to which second grade word-learning ability predicts second fourth grade vocabulary and reading comprehension after controlling for other baseline measures in second grade. Specifically, we asked:

- Does second-grade novel word learning uniquely predict fourth-grade receptive vocabulary after controlling for second-grade vocabulary, nonverbal intelligence, and SES?
- Does second-grade novel word learning uniquely predict fourth-grade reading comprehension after controlling for second-grade vocabulary, nonverbal intelligence, and SES?

3. Methodology

Data for the present study was collected between 2018 and 2023 as part of a longitudinal project examining word learning ability in typically developing children, children with dyslexia, and children with developmental language disorder. Recruitment was conducted through a combination of school-based recruitment and targeted recruitment (e.g., community advertisements) for participants with language or reading difficulties.

All procedures were reviewed and approved by the University of South Carolina Institutional Review Board. Parents provided informed consent for their children's participation, and children provided verbal assent prior to participation.

3.1. Participants

At the time of enrollment, participants (n = 172; 51% female, 48% male, 1% other) were in second grade (mean age = 8.04 years, SD = .48). The present analyses included monolingual English-speaking children who met the following inclusion criteria: (a) no history of hearing loss, vision loss, motor disorder, autism spectrum disorder, or other medical conditions as reported by parents on an intake questionnaire; (b) completion of standardized assessments of receptive vocabulary and nonverbal intelligence in second grade; and (c) completion of standardized assessments of vocabulary and reading comprehension in fourth grade.

The racial composition of the sample was White (89%), Black or African American (5.8%), Asian (6.4%), American Indian/Alaska Native (1.7%), and Other (2.3%). The majority of children were reported as Not Hispanic or Latino (95%), with 4.2% reported as Hispanic or Latino and 0.8% with unknown or unreported ethnicity. SES was indexed using eligibility for free or reduced-price school lunch as reported by parents on the intake questionnaire 19% of the sample were eligible for free or reduced-price lunch.

3.2. Standardized Measures

Participants completed standardized assessments to estimate vocabulary and cognitive abilities in second grade and evaluate children's vocabulary and reading comprehension abilities in fourth grade.

3.2.1. Peabody Picture Vocabulary Test–Fifth Edition (PPVT-5; Dunn, 2019)

Receptive vocabulary was assessed using the PPVT–5. In this task, the examiner says a target word, and the child selects one of four pictures that best corresponds to the spoken word. The PPVT-5 has strong psychometric properties, with reported internal consistency reliability of .97 and test–retest reliability of .88.

3.2.2. Test of Nonverbal Intelligence–Fourth Edition (TONI-4; Brown et al., 2010)

Nonverbal cognitive abilities were assessed using the TONI-4. The TONI-4 measures abstract reasoning and problem solving using figural matrices presented across a 3x3 grid, for which children select the best-fitting picture from a pool of 4 to 8 response choices. The test manual reports internal consistency reliabilities of .94-.96 for ages 6 to 9 years.

3.2.3. Woodcock Reading Mastery Tests–Third Edition (WRMT-3; Woodcock, 2011)

Reading comprehension was also assessed using the Passage Comprehension subtest of the WRMT-3 In this cloze-format task, children read short passages and supply missing words to complete the meaning of the text. The test manual reports internal consistencies of .93 - .96.

3.3. Word Learning Experiment

In second grade, children completed a high-context novel word learning task embedded within an interactive animated story. Participants were introduced to eight novel pseudowords as part of a fictional journey to “Kula Kula Island.” The pseudowords were constructed to conform to English phonotactic constraints.

Each novel word was paired with a visual referent (e.g., an object or animal) and presented either 18 or 35 times across the narrative. The story periodically paused to provide brief practice opportunities, during which children were asked to identify, name, or select newly learned items. Immediate feedback was provided after each response in the form of additional repetitions (e.g., “That’s a bossick!”) to reinforce learning. Following training, children completed five post-test tasks designed to assess different components of word learning: naming, listening, drawing, describing, and finding (see Table 1).

Table 1: Post-Test Tasks

Task	Instruction
Naming (phonological recall)	First, I will show you a picture and you tell me the name of it. What is this called?
Listening (phonological recognition)	I will show you a picture and you will hear four words. Listen carefully and tell me which one is the right name of the picture.
Drawing (non-verbal semantic recall)	Let's try a drawing activity. I will say a word. Then I want you to draw a picture of it on your paper.
Describing (verbal semantic recall)	Let's do a describing activity. I will say a word. Then I want you to tell me everything you know about it.
Finding (semantic recognition)	I will say a word. Then I want you to find the picture that goes with it. You can point to the picture or tell me the number.

3.4. Analyses

Analyses were conducted using participants who had complete data on second-grade predictors and each fourth-grade outcome measure. To obtain a single index of novel word learning ability, a principal components analysis (PCA) was conducted on performance across the five post-test tasks (naming, listening, drawing, describing, and finding). The first principal component accounted for 66.4% of the total variance and was retained as the composite measure of word learning ability. Higher scores on this component reflected better overall performance across the five tasks. Component scores were used as the predictor of novel word learning ability in all subsequent analyses.

Descriptive statistics were examined for all predictor variables, and regression assumptions (linearity, normality, and homoscedasticity) were evaluated and found to be reasonably met for all models. Three separate hierarchical multiple regression analyses were conducted to examine whether second-grade novel word learning predicted fourth-grade language and reading outcomes after controlling for baseline vocabulary, nonverbal intelligence, and SES. For each model, second-grade receptive vocabulary (PPVT-5), nonverbal intelligence (TONI-4), and SES (free or reduced-price lunch eligibility) were entered in Step 1. The composite novel word learning score was entered in Step 2 to evaluate its unique contribution to each outcome.

Separate models were estimated for fourth-grade receptive vocabulary (PPVT-5) and reading comprehension measured by the WRMT-3 Passage Comprehension subtest. The change in explained variance (ΔR^2) associated with the addition of novel word learning in Step 2 was used to assess its unique predictive value.

4. Results

Descriptive statistics for all examined predictor and outcome variables are provided in Table 2. Fourth grade PPVT-5 data were available for 162 children (10 missing) and WRMT-3 Passage Comprehension data for 164 children (8 missing), with minimal overlap in missingness (154 children had complete data for both outcome variables). Regression analyses predicting fourth-grade outcomes used these respective sample sizes (n = 162 for PPVT-5; n = 164 for WRMT-3).

At baseline, children demonstrated average receptive vocabulary and nonverbal intelligence scores. In fourth grade, mean receptive vocabulary and reading comprehension scores were also within the average range.

Table 2: Standardized Measures in 2nd and 4th Grade

Measure	N	Mean (M)	SD
2nd Grade Baseline Measures			
PPVT-5	172	111.61	16.33
TONI-4	172	106.79	10.74
4th Grade Outcome Measures			
PPVT-5	162	115.17	19.08
WRMT-3 Passage Comprehension	164	110.99	19.07

Pearson correlations among the predictors and outcomes are presented in Table 3. Second-grade receptive vocabulary was moderately to strongly correlated with all fourth-grade outcomes ($r = .41$ to $.61$, $p < .001$). The novel word learning composite was significantly correlated with fourth-grade receptive vocabulary ($r = .46$, $p < .001$) and passage comprehension on the WRMT-3 ($r = .40$, $p < .001$). Nonverbal intelligence was also significantly correlated with all fourth-grade outcomes, whereas SES showed weaker associations.

To examine the unique contribution of novel word learning, three hierarchical multiple regression analyses were conducted predicting fourth-grade receptive vocabulary, passage comprehension, and oral reading comprehension (Table 4).

Table 3: Pearson Correlations Among the Predictors and Outcomes

		1	2	3	4	5	6
1.	SES (School Lunch)	-					
2.	Nonverbal IQ 2nd grade (TONI-4)	-.13	-				
3.	Receptive Vocabulary in 2nd grade (PPVT-5)	-.18*	.29***	-			
4.	Novel Word Learning Composite 2nd grade	-.13	.32***	.41***	-		
5.	Receptive Vocabulary in 4th grade (PPVT-5)	-.13	.41***	.61***	.46***	-	
6.	Reading Comprehension 4th grade (WRMT-3 PC)	-.08	.40***	.47***	.40***	.55***	-

*p<.05, **p<.01, ***p<.001

4.1. Prediction of Fourth-Grade Receptive Vocabulary

In the model predicting fourth-grade receptive vocabulary, second-grade receptive vocabulary, nonverbal intelligence, and SES explained 44% of the variance in PPVT-5 scores ($F(3, 158) = 40.89$, $p < .001$) in Step 1 (see Table 4). In Step 2, the novel word learning composite was a significant predictor ($\beta = .21$, $p < .01$), accounting for an additional 3.5% of variance for a total of 47% of variance explained ($F(4, 157) = 35.01$, $p < .001$).

4.2. Prediction of Reading Comprehension

In the model predicting fourth-grade passage comprehension on the WRMT-3, second-grade receptive vocabulary, nonverbal intelligence, and SES explained 30% of the variance in WRMT-3 passage comprehension scores ($F(3, 156) = 22.43$, $p < .001$) in Step 1. In Step 2, the novel word learning composite was a significant ($\beta = .17$, $p < .05$) and explained an additional 2.2% variance in WRMT-3 passage comprehension scores. This final model explained 32% of the variance in WRMT-3 passage comprehension scores, $F(4, 155) = 18.54$, $p < .001$.

Table 4: Hierarchical Regression Models Predicting 4th Grade Outcomes

	Model Summary			Coefficients			
	R^2	ΔR^2	Sig.FΔ	b(SE)	β	t	p
Hierarchical regression predicting 4th grade receptive vocabulary (PPVT-5)							
Step 1	**.437*****						
Receptive vocabulary (PPVT-5)				.63(.07)	.54	8.64	<.001
Nonverbal Intelligence (TONI-4)				.46(.11)	.26	4.09	<.001
SES (School Lunch)				.67(2.99)	.01	0.23	.82
Step 2	**.472*****	.035	**10.22****				
Receptive vocabulary (PPVT-5)				.55(.07)	.47	7.27	<.001
Nonverbal Intelligence (TONI-4)				.39(.11)	.22	3.46	<.001
SES (School Lunch)				.98(2.90)	.02	.74	.73
Novel Word Learning Composite				2.13(.66)	.17	3.20	<.001
Hierarchical regression predicting 4th grade reading comprehension (WRMT-3 Passage Comprehension)							
Step 1	**.301*****						
Receptive vocabulary (PPVT-5)				.43(.07)	.40	5.67	<.001
Nonverbal Intelligence (TONI-4)				.45(.11)	.27	3.92	<.001
SES (School Lunch)				-.08(3.15)	0	-.03	.97
Step 2	**.323*****	.022	**5.12***				
Receptive vocabulary (PPVT-5)				.37(.08)	.34	4.60	<.001
Nonverbal Intelligence (TONI-4)				.39(.12)	.24	3.40	<.001
SES (School Lunch)				.22(3.12)	.004	.073	.94
Novel Word Learning Composite				1.62(.71)	.17	2.26	.03

***p<.05, **p<.01, ***p<.001**

5. Discussion

This study examined whether second-grade novel word learning predicts fourth-grade vocabulary and reading comprehension after controlling for second-grade receptive vocabulary, nonverbal intelligence, and socioeconomic status. Novel word learning emerged as a significant, unique predictor for both outcome variables. This pattern supports the view that experimental measures of word learning tap aspects of children's ongoing learning processes that are not fully reflected in static vocabulary scores, offering additional information about their capacity to acquire and integrate new lexical representations. These findings support the hypothesis that individual differences in word learning abilities contribute to vocabulary growth above and beyond static measures of existing vocabulary knowledge and nonverbal cognitive reasoning. Moreover, these word learning abilities are also important predictors of future reading comprehension.

These findings align with prior work that broadly highlights the importance of vocabulary to reading comprehension (Language and Reading Research Consortium & Logan, 2017; Perfetti & Stafura, 2014) and highlight the value of dynamic assessments of word learning for predicting vocabulary growth and reading comprehension. Gellert and Elbro (2013) found that Danish children's word learning efficiency assessed in third grade (age 9 years) predicted vocabulary scores measured in fourth grade, even after controlling for baseline vocabulary and nonverbal cognitive ability. The present study extends those findings by examining a larger sample of English-speaking children who were followed for two years and examining both vocabulary and reading comprehension in fourth grade. Nash et al (2025) found that English children's dynamic word learning potential assessed at 9 years of age predicted growth in vocabulary and reading comprehension measured 16 to 19 months later, even after controlling for SES, non-verbal cognitive ability, and baseline skills. Similar to the present findings, their dynamic tasks correlated less strongly with SES than static measures. Their research also showed that dynamic word learning provides unique prognostic value for the "poor comprehender" reading profile. The present findings provide initial evidence extending this work to a US context with a participant sample with wide ranging abilities; future analyses could examine word learning's prognostic value specifically for poor comprehenders within this cohort.

Dynamic assessment has been highlighted as a useful tool to inform the assessment of children's linguistic abilities in populations who have different amounts or types of language experience. In the present study, SES showed a small but significant correlation with second-grade receptive vocabulary. Importantly, word learning performance predicted outcomes even after controlling for SES, suggesting that the task captures aspects of learning efficiency that are at least partially separable from the accumulated advantages associated with socioeconomic context. However, we also acknowledge that our SES measure, which asked parents to indicate whether their child was eligible for free or reduced-price public-school lunch was offered only a limited index of SES

and had limited variability in our sample. Future work should consider using more fine-grained and multidimensional measures of socioeconomic context.

Another direction for future work concerns the measurement of reading comprehension. The current analyses focused on the WRMT-3 Passage Comprehension test, which other work has demonstrated is largely influenced by lexical knowledge (Keenan et al., 2008). While this cloze-format measure is well-suited for examining vocabulary-comprehension relationships, it does not fully capture the comprehension skills involved in reading longer, connected texts. To address this limitation, our project also included the Gray Oral Reading Tests – Fifth Edition (GORT-5; Wiederholt & Bryant, 2012), which assesses comprehension of extended passages read aloud. Future planned analyses of those data will provide a more comprehensive test of whether word learning efficiency predicts comprehension across different task formats and text lengths. The present findings have important implications for assessment. Traditional vocabulary tests provide valuable information about accumulated knowledge but are influenced by prior exposure and environmental factors. In contrast, dynamic measures of novel word learning may offer a more direct index of children's capacity to learn from instruction. Such measures may be particularly useful for identifying children who are at risk for later language and reading difficulties despite average performance on static tests, and for distinguishing children with limited exposure from those with underlying learning vulnerabilities. In addition, examining how specific components of word learning including phonological, semantic, and integration processes relate to different aspects of reading development may further clarify the mechanisms linking early word learning to later literacy outcomes.

References

Brown, Linda, Sherbenou, Ronald J., & Johnsen, Susan K. (2010). *Test of Nonverbal Intelligence: TONI-4* (4th ed.). PRO-ED.

Camilleri, Bernadette, & Botting, Nicola. (2013). Beyond static assessment of children's receptive vocabulary: The dynamic assessment of word learning (DAWL). *International Journal of Language & Communication Disorders, 48*(5), 565–581. **https://doi.org/10.1111/1460-6984.12033**

Duff, Dawna. (2019). Has vocabulary intervention had an effect? A valid, reliable, and (fairly) quick outcome measure for semantic knowledge. *Language, Speech, and Hearing Services in Schools, 50*(4), 506–517. **https://doi.org/10.1044/2019_LSHSS-VOIA-18-0134**

Dunn, Lloyd M. (2019). *Peabody Picture Vocabulary Test* (5th ed.). Pearson.

Gellert, Anna S., & Elbro, Carsten. (2013). Do experimental measures of word learning predict vocabulary development over time? A study of children from grade 3 to 4. *Learning and Individual Differences, 26*, 1–8. **https://doi.org/10.1016/j.lindif.2013.04.006**

Hadley, Elizabeth B., & Dickinson, David K. (2018). Measuring young children's word knowledge: A conceptual review. *Journal of Early Childhood Literacy, 20*(2), 223–251. **https://doi.org/10.1177/1468798417753713**

Keenan, Janice M., Betjemann, Rebecca S., & Olson, Richard K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*(3), 281–300. **https://doi.org/10.1080/10888430802132279**

Language and Reading Research Consortium, & Logan, Jessica A. R. (2017). Pressure points in reading comprehension: A quantile multiple regression analysis. *Journal of Educational Psychology, 109*(4), 451–464. **https://doi.org/10.1037/edu0000150**

Nash, Hannah, Dixon, Chris, Clarke, Paula, Oxley, Emily, Gellert, Anna S., & Weighall, Anna. (2025). Dynamic assessment of word learning as a predictor of vocabulary, reading comprehension and risk status for the poor comprehender reading profile. *Reading and Writing, 38*(9), 2633–2659. **https://doi.org/10.1007/s11145-024-10603-x**

Peña, Elizabeth D., Bedore, Lisa M., & Kester, Ellen S. (2016). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one. *International Journal of Language & Communication Disorders, 51*(2), 192–202. **https://doi.org/10.1111/1460-6984.12199**

Peña, Elizabeth, Iglesias, Aquiles, & Lidz, Carol S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology, 10*(2), 138–154. **https://doi.org/10.1044/1058-0360(2001/014)**

Petersen, Douglas B., Tonn, Penny, Spencer, Trina D., & Foster, Matthew E. (2020). The classification accuracy of a dynamic assessment of inferential word learning for bilingual English/Spanish-speaking school-age children. *Language, Speech, and Hearing Services in Schools, 51*(1), 144–164. **https://doi.org/10.1044/2019_LSHSS-18-0129**

Perfetti, Charles A., & Stafura, Joseph. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*(1), 22–37. **https://doi.org/10.1080/10888438.2013.827687**

van der Kleij, Saskia W., Burgess, Adrian P., Ricketts, Jessie, & Shapiro, Laura R. (2023). Tracking vocabulary and reading growth in children from lower and higher socioeconomic backgrounds during the transition from primary to secondary education. *Child Development, 94*(1), e57–e66. **https://doi.org/10.1111/cdev.13862**

Wiederholt, J. Lee, & Bryant, Brian R. (2012). *Gray Oral Reading Tests–Fifth Edition (GORT-5).* Pro-Ed.

Woodcock, Richard W. (2011). *Woodcock Reading Mastery Tests–Third Edition (WRMT-III).* Pearson.

Proceedings of the 50th annual Boston University Conference on Language Development

edited by Romi Hill, Xinhan Jiang, Danutham Worapipat, and Aditya Yedetore

Cascadilla Press Somerville, MA 2026

ISSN 1080-692X
ISBN 978-1-57473-047-0 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com