Research Article

Semantic–Syntactic Partial Word Knowledge Growth Through Reading

Stacy A. Wagovich,^a Margaret S. Hill,^a and Gregory F. Petroski^a

Purpose: Incidental reading provides a powerful opportunity for partial word knowledge growth in the school-age years. The extent to which children of differing language abilities can use reading experiences to glean partial knowledge of words is not well understood. The purpose of this study was to compare semantic–syntactic partial word knowledge growth of children with higher language skills (HL group; overall language standard scores of 85 or higher) to that of children with relatively lower language skills (LL group; overall receptive or expressive standard score below 85). **Method:** Thirty-two children, 16 per group, silently read stories containing unfamiliar nouns and verbs 3 times over a

1-week period. Semantic–syntactic partial word knowledge growth was assessed after each reading and 2–3 days later to assess retention.

Results: Over time, both groups showed significant partial word knowledge growth, with the HL group showing significantly more growth. In addition, both groups retained knowledge several days later.

Conclusion: Regardless of language skill level, children benefit from multiple exposures to unfamiliar words in reading in their development and retention of semantic–syntactic partial word knowledge growth.

During the school-age years, children acquire new vocabulary with impressive speed and magnitude. It has been estimated that between the third and the ninth grade, the typically developing (TD) child learns about 3,000 words per year (Nagy & Anderson, 1984; Nagy & Herman, 1987). Anglin (1993) estimated that by 10 years of age, children know close to 40,000 words. Although some of this learning stems from formal instructional activities explicitly designed to build vocabulary (e.g., Gipe & Arnold, 1979; Pany & Jenkins, 1978; Pany, Jenkins, & Schreck, 1982; Stahl, Burdige, Machuga, & Stecyk, 1992; Stahl & Fairbanks, 1986), the sheer extent of lexical acquisition has led to the widespread conclusion that most word learning occurs through incidental exposure (e.g., see Nagy & Herman, 1987, for a discussion). Children are exposed to new words through a variety of oral and written language experiences throughout their school years, perhaps most importantly through reading.

Vocabulary Growth Through Incidental Exposure in Reading

The means by which incidental exposure leads to word learning are not well understood. Carey (1978) has proposed two distinct processes: “fast mapping,” which is the knowledge a child gains from a single exposure, and “full mapping,” the more complete knowledge of a word that develops over time and with repeated exposures. According to this framework, a child’s first encounter with an unfamiliar word leads to an incomplete understanding, or partial knowledge, of that word; each subsequent experience with the word is an opportunity for additional partial word knowledge (PWK) growth. This PWK growth is hypothesized to occur across all the various domains of language, including phonological, orthographic, morpho-syntactic, and semantic. Much of the research on incidental word learning has, in fact, focused on PWK rather than complete mastery of new words. As a practical matter, Swanborn and de Glopper (1999) pointed out in their meta-analysis that learning effects from incidental exposure tended to be rather small, and studies with measures sensitive to PWK were more likely to produce greater effect sizes.

Early research on vocabulary growth through incidental reading demonstrated that TD school-age children of varying ages and ability levels can exhibit small but significant PWK gains after just a few such exposures (Konopak

^aUniversity of Missouri, Columbia

Correspondence to Stacy A. Wagovich: wagovichs@health.missouri.edu

Editor: Krista Wilkinson

Associate Editor: Carol Miller

Received March 27, 2014

Revision received July 19, 2014

Accepted October 15, 2014

DOI: 10.1044/2014_AJSLP-14-0046

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

et al., 1987; Nagy, Anderson, & Herman, 1987; Nagy, Herman, & Anderson, 1985). In fact, Wagovich and Newhoff (2004) found that even a single exposure to unfamiliar words in stories could lead to measurable gains in orthographic PWK. Some studies investigating the impact of text-related factors on the likelihood of PWK growth found that “considerate” or “elaborated” text resulted in higher rates of incidental learning (Gordon, Schumm, Coffland, & Doucette, 1992; Herman, Anderson, Pearson, & Nagy, 1987; Konopak, 1988a, 1988b); some reported that both the conceptual difficulty of the target words themselves as well as contextual factors were related to the amount of vocabulary growth (Nagy et al., 1987; Shu, Anderson, & Zhang, 1995). In contrast, Schwanenflugel, Stahl, and McFalls (1997) found that word-level characteristics, such as part of speech and concreteness, were more important than textual factors in determining the amount of PWK growth. Although some of these studies found a significant association between children’s measured language abilities and their performance on incidental learning tasks (Gordon et al., 1992; Herman et al., 1987; Konopak, 1988b), others found no such link (Konopak, 1988a; Nagy et al., 1985; Shu et al., 1995; Wagovich & Newhoff, 2004).

Incidental Word Learning in Children With Language Impairment

Taken together, these studies illustrate how TD children can, in fact, learn new vocabulary incidentally through silent reading. Although little research to date has examined incidental word learning through silent reading in children with language impairment, several studies of “quick incidental learning” (QUIL; Rice & Woodsmall, 1988) through *oral* exposures to words have been conducted with young children with specific language impairment (SLI). The QUIL paradigm involves presenting children with a brief video along with an oral narrative that introduces new (often novel) words incidentally. Results from several of these studies indicate that children with SLI do exhibit partial word learning in the QUIL context, but their performance lags behind that of their TD peers (e.g., Horohov & Oetting, 2004; Oetting, Rice, & Swank, 1995; Rice, Oetting, Marquis, Bode, & Pae, 1994).

One study that examined partial word learning through silent reading in children with language impairment yielded a similar pattern of results (Steele & Watkins, 2010). The authors compared the PWK growth of fourth- and fifth-grade children with language-learning disability (LLD) to that of their TD peers. All of the children in the LLD group had received a diagnosis of language impairment (either expressive only or mixed), and some had been diagnosed with a reading comprehension disability as well. The children silently read stories containing novel words, followed by oral and written tasks assessing syntactic and semantic PWK growth. Both groups demonstrated learning of the novel words, but the LLD group’s performance was significantly poorer than that of the TD group. Importantly, all of the children were made aware before they read the stories

that the passages would include some nonsense words and that part of their task was to infer the meanings of these words. Thus, even in relatively supportive conditions, which included alerting the children to the words in advance, the LLD group’s learning fell significantly short of that of the TD children.

There is also evidence that children with language weaknesses, who do not meet traditional score cutoffs for clinical language impairment, may exhibit subtle deficits in incidental word learning. A recent study by Wagovich, Pak, and Miller (2012) compared growth in orthographic PWK through silent reading among 10- to 16-year-old children with lower language skills (LL group) to that of age- and gender-matched peers with higher language skills (HL group). Participants in the LL group scored at least 1.0 standard deviation below the mean on receptive or expressive language. Results showed that both groups demonstrated learning of the orthographic forms of the experimental rare words over time, and differences between groups were not significant. However, the children in the LL group incorrectly identified orthographically similar nonword foils as real words significantly more often than the HL group, suggesting that the orthographic representations of these children may not have been as refined as those of their typically developing peers.

The Effect of Repetition on Incidental Word Learning

A second focus of our study concerned the effect of repetition on PWK growth. We might expect that given a greater number of incidental exposures to unfamiliar words, children will demonstrate greater gains in PWK. There may be important differences, however, depending on the conditions in which the repetition occurs. In a study by Horst, Parsons, and Bryan (2011), 3-year-old children heard stories containing novel word–object pairs on three separate occasions over the course of a week. Half of the children were exposed to word–object pairs in the same stories, and the other half were exposed to the pairs in different stories. Posttesting following each of the three sessions showed that children exposed to novel word–object pairs from the same story, read three times in a row, learned more words than children who heard the same words from three different stories. Children who heard the same stories also had better retention; on average, they recalled two thirds of the words, whereas children who heard different stories performed no better than chance. The authors speculated that because of the contextual repetition afforded by hearing the same story three times in a row, the children in the same-stories group were free to devote more of their attentional resources to the novel aspects of the stories, the novel word–object pairs themselves.

QUIL studies of young children have shown that increasing the number of learning trials results in enhanced performance; however, the effect may not be the same for children with SLI as compared to TD children. Rice et al. (1994), for example, found that TD 5-year-old children made

significant gains in vocabulary with only three repetitions, whereas the SLI group required 10 exposures to achieve comparable scores. In an oral incidental learning paradigm, Nash and Donaldson (2005) found that for semantic PWK, 5- to 9-year-old children with SLI continued to perform more poorly after 12 exposures than their age-matched peers did after six exposures. In the context of silent reading, whether a similar pattern is found among older, school-age children with language impairment, or even among children with weaker language relative to peers, is a question with significant theoretical and practical implications. These children may require more direct instruction and more repetition than peers with stronger language skills in order to acquire new vocabulary.

Studies examining the effect of repetition on PWK growth through reading have yielded conflicting results. Schwanenflugel et al. (1997) found no difference in performance on semantic PWK tasks after one versus five exposures to rare words in stories among TD children; however, most of the words were presented only once. Similarly, Steele and Watkins (2010) reported no significant difference in semantic or syntactic PWK growth, given two versus five exposures, for either LLD or TD children. In contrast, Wagovich et al. (2012) found that, for children with both lower language and higher language, orthographic PWK of rare words increased over successive trials as the children read stories three times. There is surprisingly little evidence to date that children with language impairment, or lower language, respond differently than children with higher language to the number of exposures through reading. In a read-aloud task with 9- to 10-year-olds, Cain, Oakhill, and Lemmon (2004) did report that children with lower comprehension and vocabulary skills required significantly more trials to earn comparable scores to children with higher skills. Importantly, however, it appears that no recent investigation has found a difference in PWK growth between children of lower and higher language levels in their response to repetition in *silent* reading contexts.

Part of Speech and PWK Growth

Another focus of our study was on whether children with higher language or lower language differ in their pattern of incidental learning of nouns as compared to verbs. Until recently, most studies examining this question have been oral fast-mapping or QUIL studies conducted with younger children. Some of this research indicates that nouns are easier to learn than verbs. In a fast-mapping task with 4- to 6-year-old children, Alt, Plante, and Creusere (2004) reported that both SLI and TD groups learned more object than action labels from context, and they were able to map more semantic features onto objects than onto actions. In a QUIL study of 6- to 8-year-old children that included words from four semantic classes (object, attribute, action, and affective state), Oetting et al. (1995) found that both SLI and TD groups had their greatest gains on object words. In contrast, Horohov and Oetting (2004) found that scores for verbs were significantly higher than for nouns among young children

with SLI as well as both age-matched and language-matched controls. In a similar vein, Rice et al. (1994) reported that after 10 repetitions, both TD and SLI groups learned more verbs than nouns. Clearly, it is not a simple case of nouns being easier to learn from oral exposures than verbs, either for TD children or those with SLI.

Few studies involving reading tasks with school-age children have examined the issue of grammatical class and its impact on the likelihood of incidental word learning; results so far have been conflicting. Steele and Watkins (2010) reported a trend for TD and LLD groups to perform better with nouns than with verbs, but the difference was not significant. Schwanenflugel et al. (1997), on the other hand, found that non-nouns were acquired more often than nouns, and Wagovich and Newhoff's (2004) participants showed significant orthographic PWK growth for verbs but not for nouns. The role of children's syntactic understanding in the course of vocabulary growth has both theoretical and practical implications. From a theoretical standpoint, syntactic and semantic PWK are interconnected because knowledge of a word's grammatical class is related to better understanding of semantic relations (e.g., agent, action, object) and therefore adds to a child's understanding of that word's meaning. From a practical standpoint, including a measure of syntactic PWK allows for greater sensitivity in the evaluation of overall word knowledge growth.

Retention of Incidental Word Learning

Presumably, if the PWK gained incidentally is to accumulate and grow, over time and with repeated encounters, into more complete representations, children must also be able to retain the PWK from these incidental exposures. There is evidence that even very young TD children demonstrate such retention after shared book reading experiences, as reported by Horst et al. (2011). QUIL studies of young children with SLI, on the other hand, suggest that there may be particular problems with retention, especially with verbs (Oetting, 1999; Rice et al., 1994). In a silent reading context, Wagovich et al. (2012) reported good retention of orthographic PWK among both lower and higher language groups after a delay of 2–3 days. However, orthographic PWK may present a lesser linguistic challenge than the formation of syntactic or semantic representations. Indeed, several studies have shown that, at least among TD children, learning of the orthographic forms of words can occur quickly, with few exposures (e.g., Bowey & Muller, 2005; Cunningham, Perry, Stanovich, & Share, 2002; Nation, Angell, & Castles, 2007; Share, 1999). To what extent children retain semantic–syntactic PWK gained in the silent reading context, and whether this process differs for children of varying ability levels, is an important question that has received scant research attention.

In summary, incidental PWK gained through silent reading is a major source of vocabulary growth during the school-age years. However, the extent to which children of varying language abilities are able to accomplish word learning through incidental reading exposure is not well

understood. In addition, there may be effects of repetition and part of speech that interact with language skill and affect the likelihood that children will gain vocabulary knowledge during silent reading. Moreover, very little is known about how well school-age children of varying language skill levels actually retain the PWK they have achieved through their silent reading experiences.

Purpose of the Study

We undertook this study to compare children with higher language to children with relatively lower language in their ability to glean PWK from encountering unfamiliar words multiple times in reading. In particular, we asked the following questions:

- Do children with higher language versus lower language demonstrate PWK over the course of multiple exposures to unfamiliar words in reading? If so, do the groups differ in their pattern of learning?
- Do children retain PWK gains 2–3 days following their final exposure to the words?
- Is there a difference in the patterns of noun versus verb learning, and do the groups differ in these patterns?

Method

Participants

Sixteen children with higher language (HL group) and 16 children with relatively lower language (LL group) participated in the study. Children in the LL group were pair-matched to children in the HL group based on age and gender. The children ranged in age from 10;6 (years;months) to 16;5. The average age difference within age-matched pairs was 2.6 months ($SD = 2.3$). Each group consisted of four males and 12 females. Fourteen of the children in each group were the same as those in Wagovich et al. (2012).¹ Children were recruited through flyers and electronic announcements posted throughout mid-Missouri. All children were monolingual English speakers. To participate, all passed a bilateral hearing screening. All were reported by parents to have normal vision, corrected or uncorrected, and no history of neurological impairment, either developmental or acquired. In addition, all participants passed a screening of nonverbal reasoning, the Symbolic Relations subtest of the Detroit Test of Learning Aptitude–Fourth Edition (DTLA-4; Hammill, 1998), receiving a subtest standard score of 7 or higher ($Test M = 10$, $SD = 3$).

To be included in the LL group, children had to receive a standard score lower than 85, performing more than 1.0 standard deviation below the mean on the receptive or expressive portion of the Clinical Evaluation of Language

Fundamentals–Fourth Edition (CELF-4; Semel, Wiig, & Secord, 2003). As discussed in Wagovich et al. (2012), 18 children were excluded from the LL group because, despite parental concerns about language, CELF-4 scores were too high for inclusion in this group. It should be noted that the children in the LL group did not all meet the traditional criterion for language impairment of -1.25 standard deviations of the mean (with standard scores of 81 or lower) on receptive/expressive language. Rather, approximately 69% (11/16 children) received scores of 81 or lower on the receptive composite, the expressive composite, or both.

Receptive vocabulary standard scores were permitted to vary within the LL group, with standard scores on the Peabody Picture Vocabulary Test–Third Edition (PPVT-3; Dunn & Dunn, 1997) ranging from 74 to 122. Only one child scored within the traditional clinical range of -1.25 standard deviations of the mean (i.e., receiving a standard score of 81 or lower). In addition, reading standard scores, as measured by the reading subscales of the Woodcock–Johnson III Tests of Achievement (WJ-3; Woodcock, McGrew, & Mather, 2001), were permitted to vary for the LL group. Children received the Word Identification, Word Attack, and Passage Comprehension subscales of the WJ-3. On Word Identification, two children received a standard score of 81 or lower, and on Word Attack and Passage Comprehension, none of the children received scores of 81 or lower. Thus, we describe the LL group not as a “language impairment” group but as displaying lower overall language skills relative to the HL group.

To be included in the HL group, children had to receive a standard score of 85 or higher on the receptive and expressive portions of the CELF-4. In addition, all of the children in this group received scores of 85 or higher on the PPVT-3 and the WJ-3 reading subscales. Children in this group had no history of language or learning difficulties and were reported by parents to be progressing satisfactorily in school. In addition, parents of this group reported no concerns about their children’s language or learning. Table 1 summarizes the language test results for each group.

Parents provided descriptive information about maternal education. As discussed in Wagovich et al. (2012), it was not feasible to pair-match; however, maternal education was fairly similar between groups. For the LL group, 75% of the mothers had some college (33% of those being college graduates), compared to 94% of the mothers in the HL group (93% of those being college graduates). Both groups had one parent reporting maternal education as non-high school graduate. Thus, the majority of both groups of mothers had received education at the college level.

Stimuli

Rare words. Thirty-two rare words were selected for the study. The words were the same as those described in Wagovich et al. (2012). The words were selected by consulting several rare word sources. All of the words were sufficiently rare that they were not included as an entry in Carroll, Davies, and Richman’s (1971) word frequency data set.

¹Two children in the LL group and their matched pairs from the previous study were excluded from this study because the LL group children received an earlier version of the multiple choice posttest, the measure of interest in the present study. These pairs were replaced by two additional pairs that met all criteria for inclusion in this study.

Table 1. Standardized test scores for lower language skills (LL) and higher language skills (HL) groups.

Measure	LL group	HL group
PPVT-3		
<i>M</i>	95.9**	119.9
<i>SD</i>	11.6	8.7
<i>Mdn</i>	96	118
Range	74–122	105–136
CELF-4		
Receptive Composite		
<i>M</i>	83.4**	111.4
<i>SD</i>	7.0	10.5
<i>Mdn</i>	82.5	113
Range	68–99	96–131
Expressive Composite		
<i>M</i>	84.8**	110.4
<i>SD</i>	9.7	8.6
<i>Mdn</i>	86	111
Range	63–99	98–124
WJ-3		
Word Identification		
<i>M</i>	91.2**	107.9
<i>SD</i>	5.9	9.9
<i>Mdn</i>	93	109
Range	79–99	91–126
Word Attack		
<i>M</i>	92.6*	101.7
<i>SD</i>	5.6	10.5
<i>Mdn</i>	92	99.5
Range	86–104	86–126
Passage Comprehension		
<i>M</i>	94.1**	109.1
<i>SD</i>	5.5	8.1
<i>Mdn</i>	93	107
Range	87–104	100–129

Note. Standard scores *M* = 100, *SD* = 15. PPVT-3 = Peabody Picture Vocabulary Test–Third Edition (Dunn & Dunn, 1997); CELF-4 = Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel et al., 2003); WJ-3 = Woodcock–Johnson III Tests of Achievement (Woodcock et al., 2001).

Independent-samples *t* test revealed significant between group differences: **p* < .01. ***p* < .001.

In addition, each word was two syllables in length and was consistent with the content of the story in which the word was to be inserted. Half of the words were nouns, and half were verbs. A list of the words is provided in Appendix A.

Stories. The four stories in which the rare words were to be embedded were selected from sixth-grade curriculum readers. As described in Wagovich et al. (2012), the stories were modified slightly so that they were similar in length (2,569 words to 2,979 words) and their readability indices were similar (within the fourth-grade level according to Flesch–Kincaid grade-level scores; Flesch, 1974). The 32 rare words were inserted once each into the four stories such that each story contained single occurrences of four rare nouns and four rare verbs. Stories were printed on white paper with 1.5-point line spacing and using a 14-point font. Stories ranged in page length from 12 to 15 pages. A complete description of the stories is provided in Wagovich and Newhoff (2004).

Multiple choice posttest. The multiple choice measure included a set of questions that focused on the syntactic category, followed by a set of questions that focused on the general semantic domain of the rare words. Sample items are provided in Appendix B. The syntactic category questions were identical, with each question asking whether the word was (a) a “person, place, thing, or idea,” or (b) an “action word.” A third, “don’t know,” option was provided, but there was not a “none of the above” option. Therefore, the test taker was intended to infer that each item was either a noun or a verb. The “don’t know” option was intended to discourage random guessing of (a) or (b).

The general semantic domain questions were intended to measure whether the test-taker knew general category information about the rare words. Fifteen categories (e.g., “clothing or something worn,” “movement,” “a state or a feeling”) were developed that encompassed all 32 rare words, and these categories were used to construct the answer choices. To the extent possible, the frequency with which each category served as an answer choice was controlled so that the presence of a less frequently occurring answer choice would not cue the test-taker as to the right answer. Each category was used as an answer choice 10–13 times within the test. Once again, a “don’t know” option was included to discourage random guessing. A “none of the above” option was not included, because it was important for test-takers to infer that each question had a correct answer listed. This was particularly important for this section of the test because words can be members of several semantic domains, only one of which would be represented in a question. Knowing that the right answer was definitely listed as an answer choice would enable children to think of each word in light of the choices provided and try to find a good match between their understanding of the word and the choices presented. Both sections of the multiple choice posttest included questions for 12 common words (e.g., *forest*, *highway*, *hurry*, *cancel*). The common words were taken from the stories and were sufficiently frequent within the English language that the children in the study would know their meaning and have encountered the words in other reading prior to the study.

Three randomized versions of the posttest were developed. Over the four posttest administrations per child, each child received each version at least once. The order of administration was randomized across participants. Thus, any fatigue effects in completing the 44-item test should have been evenly distributed across the rare word items.

Procedure

Participants who completed the study attended a total of six individually administered sessions. All children attended two initial testing sessions to complete the language and reading testing as well as the nonverbal reasoning and hearing screenings. In the first of these two sessions, the children completed a pretest, detailed in Wagovich et al. (2012), in which they were asked to circle words if they looked at all familiar and to define or write a sentence for any words

they knew. For the present study, this pretest was used to evaluate whether the children knew the rare words selected for inclusion in the stories.

Children who met inclusion criteria for one of the two groups attended four additional sessions in which the experiment was administered. In the third, fourth, and fifth sessions of the study, each 2–3 days apart, children read the same two stories (of four possible, counterbalanced). Children sat with an examiner and were instructed to take their time and read the stories silently. The examiner observed the child while reading but did not provide any directives or cues as the child read. The reading experience was intended to be similar to the natural reading experiences of the child. After reading each story, children were asked to summarize it orally. Next, children were given a sheet of paper with a maze printed on the front. They were told that they were taking a break from the experiment and to work on the maze for a few minutes. They were asked to let the examiner know when they were ready to move on to the next part of the session. Most children spent 3–5 min working on the maze. Next, the multiple choice posttest was administered to assess PWK growth that occurred from reading the stories. Thus, by Session 5 of the study, participants had read two stories three times each and had completed PWK posttests after each reading. In Session 6, occurring 2–3 days later, the children completed only the posttesting in order to assess their retention of the rare words encountered in the stories they read. The four sessions in which the posttest was administered are referred to hereafter as Time 1 (T1), Time 2 (T2), Time 3 (T3), and Time 4 (T4)/Retention.

Computing PWK Scores

In scoring the multiple choice posttest, rare word and common word scores were computed separately. For each word, one point was awarded for a correct response to the syntactic item, and one point for a correct response to the general semantic item, so that scores for each word ranged from 0 to 2. Common words were analyzed first to determine whether each child was reasonably accurate in responding to known words for each of the four posttests. We reasoned that if a child was not accurate in responding to known words on a particular posttest, the child either did not have an adequate understanding of the task or was not attending to the task sufficiently, thus invalidating the child's responses to the rare words. Data were then removed from the data set for any child's session if the child did not complete the common word items with at least 70% accuracy for that session. This procedure resulted in the removal of one participant from the LL group entirely, because performance on the common words did not reach the criterion at any of the four time points. Because this participant was removed from analyses entirely, his matched control in the HL group was removed as well. In addition, one LL group participant did not reach criterion for the common words at T2 or T4, and an additional LL group participant did not reach this criterion for T2, T3, or T4. The data for these

two children were only removed for the sessions in which they did not reach the 70% criterion for known words.

PWK for the rare words was assessed by computing the posttest score on the rare words that a child encountered in the two stories he or she read (experimental words; eight nouns and eight verbs). A posttest score was also computed for the rare words in the stories that the child did not read (control words; eight nouns and eight verbs). Thus, the control word score was taken into account in evaluating the experimental word learning that occurred. This procedure is viewed as more stringent than comparing individual experimental word scores to chance effects.

Analyses

A repeated measures, mixed model analysis of covariance (ANCOVA) was used, with experimental word learning score as the outcome variable, group (LL, HL) as the between-subjects factor, and time (T1, T2, T3, T4/Retention) and part of speech (noun, verb) as within-subjects repeated factors. The analysis of variance portion of the model included the main effects plus all two-way and three-way interaction terms. Control word learning was the time-varying covariate and thus was not included in the interaction effects. For this analysis, the covariance structure for part of speech was unstructured, and compound symmetry was used for the test factor. "Subject" was treated as a random effect. The choice of covariance structures was guided by minimizing the Akaike Information Criterion (Akaike, 1974). Denominator degrees of freedom are based on the Kenward–Roger method (Kenward & Roger, 1997). Histograms and normal probability plots were used to examine the distribution of the residuals. The "mixed" procedure in SAS/STAT (Version 12.1; SAS Institute, 2012) was used for the analysis.

Results

As can be seen in Table 1, the groups differed on all language and reading measures, with the LL group scoring significantly lower than the HL group. These differences were expected, given that language score on the CELF-4 was the primary grouping variable. Despite group differences on the three WJ-3 reading subscales, it should be noted that none of the children, with the exception of 2 LL group members on Word Identification only, received scores of 81 or lower on the reading measures. In addition, the groups did not differ in nonverbal reasoning, as measured by the Symbolic Relations subtest of the DTLA-4.

As an initial step, we examined the children's pretest checklists to determine whether they provided sentences or definitions that suggested some prior knowledge of the rare words. Of the children in the LL group, none were able to provide a meaningful sentence or definition for any of the rare words in the study; of the HL group, two children wrote meaningful sentences or definitions for a single rare word each, and the rest of the children in the HL group wrote none. In comparison, for the common words in the checklist, the LL group provided a meaningful sentence or

definition 93.9% ($SD = 9.2$) of the time, and the HL group provided one 95.6% ($SD = 8.8$) of the time.

PWK

Prior to model interpretation, an analysis of residuals revealed a unimodal, light-tailed distribution with good fit to the normal distribution. Table 2 gives the full results for the analysis of covariance model. The statistically significant factors include main effects for Group ($p = .006$) and Time ($p < .0001$). The main effect for Part of Speech was not significant ($p = .88$). There were no significant interaction effects for Group $\times$ Time ($p = .12$), Group $\times$ Part of Speech ($p = .12$), or Group $\times$ Time $\times$ Part of Speech ($p = .38$).

Figure 1 displays the mean covariate-adjusted experimental word scores for each group over time. Descriptively, the pattern of means differs by language group. Indeed, pairwise comparisons reveal that for the HL group, the T1 to T2 difference is significant ($p = .03$), whereas for the LL group, the T2 to T3 difference is significant ($p = .04$). With respect to retention of word knowledge, there was no significant difference for either group from T3 to retention testing at T4 (LL group: $p = .09$; HL group: $p = .96$).

As indicated above, the data depicted in Figure 1 are derived word learning scores, expressing the amount of experimental PWK while controlling for the control PWK. Table 3 provides the raw scores for each group on the experimental and control words for each session, as a supplement to the derived scores on which Figure 1 is based.

Discussion

The focus of this study was on the development of semantic and syntactic PWK over the course of multiple exposures to stories containing unfamiliar words over time. School-age children with relatively lower language (LL group) and those with higher language (HL group) read stories containing half of the unfamiliar words, presented one time each, and the children's syntactic and semantic PWK were measured after each reading. The amount of PWK growth, over time and between groups, was analyzed while controlling statistically for the PWK that occurred on the half of the words not encountered in the stories (control words). The two main findings were that (a) overall, the children

in both groups demonstrated significant PWK growth over time, but that (b) the LL group demonstrated significantly less PWK growth overall than the HL group. In addition, both groups retained the PWK several days later.

An earlier study by Wagovich and Newhoff (2004) directly informed this work. The study examined the development of PWK in TD children, given one exposure to a set of unfamiliar nouns and verbs embedded in stories. The main finding as relates to this study was that one exposure was not sufficient to detect either the children's semantic or their syntactic PWK growth. The present study addressed this issue in two ways. First, our study was designed with three exposures to the stories containing the words, with PWK measurement after each exposure. Second, in assessing word knowledge, we attempted to enhance measurement sensitivity by merging syntactic and semantic item types into one PWK score (see Appendix B). The reasons for this were both practical and theoretical. From a practical perspective, merging the item types creates a more robust measure of PWK. From a theoretical perspective, syntactic and semantic knowledge are interrelated, in that knowledge of a word's syntactic category as noun or verb is related to knowledge of semantic class (e.g., agent, object, action). This knowledge can be accompanied by some additional meaning information. For example, burgeoning understanding of a word as a verb and an action may be accompanied by basic information about the nature of the action denoted by the word.

Our first main finding is that, overall, children in both groups demonstrated PWK growth over time given a sufficient number of exposures to the stories containing the words. These findings suggest that even children with relatively lower language can acquire (or begin to acquire) new words from their incidental reading experiences. As can be seen from Figure 1, both groups displayed an overall pattern of growth over time.

Importantly, the statistical model employed estimates growth from the reading condition alone, controlling for PWK growth that occurred on words not encountered in the stories (control words). However, PWK growth clearly occurred on these control words. This is to be expected when the items are real lexical items, as opposed to nonwords; children are able to use their knowledge of morphological, phonological, and orthographic regularities to infer a word's

Table 2. ANCOVA results for lower language skills and higher language skills groups over time.

Effect	Numerator <i>df</i>	Denominator <i>df</i>	<i>F</i>	<i>p</i>
Group	1	26.6	8.99	.0058
Time	3	79.2	9.59	<.0001
Group $\times$ Time	3	79.5	1.99	.1218
Part of speech	1	26.9	0.02	.8760
Group $\times$ Part of Speech	1	28.4	2.63	.1157
Time $\times$ Part of Speech	3	77.0	0.68	.5653
Group $\times$ Time $\times$ Part of Speech	3	76.9	1.04	.3786
Control word score ($\beta = .237$)	1	207.0	10.07	.0017

Note. ANCOVA = analysis of covariance.

Figure 1. Model-implied means for each lower language skills (LL) and higher language skills (HL) group over time (error bars = 95% confidence interval). T1 = Time 1; T2 = Time 2; T3 = Time 3; T4 = Time 4.

syntactic and semantic information (e.g., White, Power, & White, 1989). The word *adduce* is an example of this; orthographic information could enable a child to infer correctly that the word is a verb, and the prefix could enable the meaning inference that the word has something to do with “adding.” The results suggest that children in both groups were using linguistic knowledge to infer new information about all words, both those encountered in the stories and those encountered only in the assessment phases of the experiment. Importantly, the study was designed to control for the natural PWK growth that occurs from simply seeing words in isolation and deriving some meaning using linguistic knowledge. The PWK growth detected through our statistical procedure and displayed in Figure 1 reflects that

which occurred through reading unfamiliar words in context, while taking into account PWK of control words.

As might be expected, a second finding is that the LL group demonstrated significantly less overall PWK growth than the HL group. This finding differs from Wagovich and Newhoff’s (2004) finding of no significant relationship between language ability and PWK growth. The design difference that likely accounts for the difference in findings is that the present study includes children with a wider range of language skills; the LL group all had receptive or expressive language scores at least one standard deviation below the mean. Although our LL group, on average, demonstrated stronger language skills than those of children within the SLI literature, it is noteworthy that our findings are

Table 3. PWK of experimental and control words: means, standard deviations, and ranges for raw data.

Session	LL group				HL group			
	Experimental word score		Control word score		Experimental word score		Control word score	
	<i>M</i> (<i>SD</i>)	Range	<i>M</i> (<i>SD</i>)	Range	<i>M</i> (<i>SD</i>)	Range	<i>M</i> (<i>SD</i>)	Range
Time 1								
Nouns	5.5 (3.5)	1–12	5.2 (2.6)	1–9	7.9 (2.1)	5–12	7.4 (2.9)	1–13
Verbs	5.5 (1.9)	3–10	5.9 (2.1)	2–10	6.8 (3.0)	2–11	5.5 (2.0)	2–8
Time 2								
Nouns	4.7 (3.0)	1–11	5.2 (2.5)	1–9	8.5 (2.5)	4–13	7.6 (1.2)	6–9
Verbs	5.6 (3.0)	0–11	6.2 (2.5)	3–11	8.3 (3.0)	4–13	6.6 (2.0)	3–9
Time 3								
Nouns	5.6 (3.0)	2–13	5.7 (2.9)	1–10	9.7 (2.8)	6–15	8.1 (1.8)	4–11
Verbs	6.9 (2.2)	3–10	5.9 (2.1)	3–10	8.4 (2.7)	4–14	6.7 (1.5)	3–9
Time 4 (Retention)								
Nouns	6.4 (3.9)	1–15	5.2 (3.1)	0–11	9.9 (3.2)	4–14	8.3 (2.7)	3–12
Verbs	7.2 (2.8)	1–11	6.0 (2.6)	1–10	8.4 (3.0)	4–14	7.1 (1.8)	4–10

Note. Scores are based on a maximum possible score of 16 (8 words × 2 points each). PWK = partial word knowledge; LL = lower language skills; HL = higher language skills.

consistent with research comparing the semantic–syntactic PWK growth of children with SLI to their TD peers on reading tasks (Steele & Watkins, 2010). Steele and Watkins (2010) found a significant word learning effect for language ability. It is worth noting that if the children in the LL group of this study had performed within the clinical range for SLI, the effect of group may have been greater.

We do not wish to overinterpret our group contrasts across individual time periods, given an interaction that was not significant. However, it is worth noting that whereas the HL group showed a significant increase in PWK from T1 to T2 (following two exposures to the stories), the LL group required three exposures before demonstrating a significant increase. It should be noted, of course, that the groups showed considerable variability throughout the study, as can be seen in the large confidence intervals in Figure 1.

A key finding was that of no significant main effect for part of speech. Overall, there was no difference between groups or over time in the children's PWK of nouns compared to verbs. For this analysis, it is also clear that the groups displayed considerable variability in performance. Our findings are somewhat in contrast to those of other similar studies. Wagovich and Newhoff (2004) found greater word discrimination PWK growth on verbs than on nouns, and Schwanenflugel et al. (1997) reported greater semantic learning for non-nouns than for nouns. However, our results are consistent with those of Steele and Watkins (2010), whose study also focused on syntactic and semantic PWK. As discussed, their study included a group of children with language impairment, unlike the studies by Wagovich and Newhoff and Schwanenflugel et al. Although the present study did not include a language impairment group, it included children representing a fuller range of language skills than the studies by Wagovich and Newhoff and Schwanenflugel et al. In this light, it is not surprising that our findings with respect to part of speech are more similar to those of Steele and Watkins, who found no overall significant main effect for part of speech.

Our finding that both groups of children did retain the PWK acquired through reading is encouraging and somewhat thought-provoking. As can be seen from Figure 1, the HL group mean remains close to the same from T3 to the retention phase at T4. However, the LL group mean actually increased from T3 to the retention phase (this increase approached significance). Because the children were not exposed to the stories at T4, additional learning is not likely. There are at least two possible explanations. First, it is possible that, following the three readings and PWK posttests, the children in the LL group were able to reflect on their encounters with the words in a substantive way that was not possible after only two readings (recall that this group demonstrated significant PWK growth from T2 to T3). It is possible that such reflection resulted in better performance at Retention than at T3. A second, perhaps more likely, explanation is that at T4, because there were no stories to read, children in the LL group were less fatigued than in previous sessions and better able to devote

their full attention and effort to the assessment. In support of this possibility is the informal observation that the children in the LL group generally took longer to read the stories, making the total length of Sessions 1, 2, and 3 longer for this group. If this is the case, then it is likely that the LL group's performance at T1, T2, and T3 was somewhat artificially depressed and that the best estimate of their PWK growth over the course of the study is their performance at T4, in the retention phase.

Study Limitations and Future Directions

Several limitations of the study should be noted. First, although most children with language impairment are males, this study consisted of more female than male participants. This could impact the generalizability of findings. Second, rare words in the English language were used as stimuli for the study. Although the words were piloted with this age group in order to ensure that the words would be unfamiliar to the children in our study, it is possible that the children may have seen them before. Using nonwords with contrived meanings would have solved these issues but likely created other study design problems. As Nagy et al. (1985) have suggested, developing a novel word and a meaning corresponding to it will often lead to the creation of a meaning that is merely a synonym of a word children know, without the semantic complexity of an actual word within the language. Moreover, for older children, the insertion of nonwords within a story may be obvious and, thus, cause the child to focus more on the nonword vocabulary targets than on the content of the story.

In addition, from this study, it is not possible to determine the extent to which learning occurred following the first exposure. To measure this would have required administration of the multiple choice task prior to the children encountering the words in the stories for the first time. Given that the multiple choice task (unlike the pretest checklist) required thinking about each of the words and their meanings relative to the answer choices, we opted not to give this task to the children first, so that the reading experience would more closely resemble incidental word encounters. However, we suspect that only minimal word learning occurred from the first exposure for two reasons. First, Wagovich and Newhoff (2004) administered a multiple choice task very similar to the present task to a group of children similar to our HL group and found no significant syntactic or general semantic word knowledge growth from one exposure alone. Therefore, it is likely that the children in the present study performed similarly. Second, as Table 3 suggests, the amount of experimental PWK demonstrated following the first exposure in the stories is similar to the amount of control PWK (across nouns and verbs, LL group: 5.5% vs. 5.6%; HL group: 7.4% vs. 6.5%, respectively), indicating that the amount of PWK growth attributable to the story presentation was fairly minimal. In future investigations, perhaps a "pretest multiple choice task" could be administered well in advance of the study (1–2 weeks). This would enable comparisons from pretest to T1 while having

less of an impact on participants in terms of their focusing on the words instead of the stories.

Finally, as discussed above, measurement of PWK in the LL group was likely impacted by test fatigue in Sessions 1, 2, and 3, in which the children had to read two stories and then complete the PWK assessment. We attempted to minimize test fatigue by asking the children to complete a maze (intended as a fun, nonverbal diversion) prior to beginning the PWK posttest. Perhaps in future studies, children could be asked to leave the building and return in an hour (i.e., require a more substantial break of the children). Another tack could be to reduce reading demands by choosing shorter stories. However, there is evidence that stories with high concentrations of unknown words negatively impact word learning (Swanborn & de Glopper, 1999). It should be noted that even though test fatigue in the LL group is a strong possibility, the children in this group still demonstrated significant PWK growth over the course of the study, given the language and reading demands of the study.

Clinical Implications

The results of this study are encouraging from a clinical perspective. The study was undertaken to determine whether children with relatively lower language could benefit, as their peers with higher language do, from incidental reading to acquire new vocabulary. The results of the study suggest that, indeed, even children with lower language benefit from these informal reading experiences and the exposures they provide to new words. Importantly, the study is consistent with the literature in this area in that it provides some evidence that multiple exposures may be needed for the instantiation of new vocabulary, particularly for children for whom language skills are relatively weaker. Moreover, our findings indicate that, given several exposures to the words in the same contexts over time, children with lower language, as well as their peers with higher language, seem to retain the vocabulary knowledge at least for 2–3 days following exposure.

In interpreting findings of this study relative to the clinical population of children with language impairments, it is important to note that children with language impairments may need even more written exposures to a word and/or additional learning supports to instantiate new word knowledge and retain that knowledge. That is, knowledge growth may occur in even smaller increments than was observed in our LL group. This underscores the importance of using sensitive measures of PWK to detect learning.

Acknowledgments

This work was funded by National Institute on Deafness and Other Communication Disorders Grant R03DC006827-01A1, awarded to the first author. We gratefully acknowledge the assistance of the following students: Christine Franke, Janine Ingram, and Chesney Willis.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.
- Alt, M., Plante, E., & Creusere, M. (2004). Semantic features in fast-mapping: Performance of preschoolers with specific language impairment versus preschoolers with normal language. *Journal of Speech, Language, and Hearing Research*, 47, 407–420.
- Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development*, 58(10, Serial No. 238).
- Bowey, J. A., & Muller, D. (2005). Phonological recoding and rapid orthographic learning in third-graders' silent reading: A critical test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 92, 203–219.
- Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, 96, 671–681.
- Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.
- Carroll, J. B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. Boston, MA: Houghton Mifflin.
- Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: Examining the role of self-teaching. *Journal of Experimental Child Psychology*, 82, 185–199.
- Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test—Third Edition*. Circle Pines, MN: AGS.
- Flesch, R. F. (1974). *The art of readable writing with the Flesch readability formula*. New York, NY: HarperCollins.
- Gipe, J. P., & Arnold, R. D. (1979). Teaching vocabulary through familiar associations and contexts. *Journal of Reading Behavior*, 11, 281–285.
- Gordon, J., Schumm, J. S., Coffland, C., & Doucette, M. (1992). Effects of inconsiderate versus considerate text on elementary students' vocabulary learning. *Reading Psychology: An International Quarterly*, 13, 157–169.
- Hammill, D. D. (1998). *Detroit Tests of Learning Aptitude—Fourth Edition*. Austin, TX: Pro-Ed.
- Herman, P. A., Anderson, R. C., Pearson, P. D., & Nagy, W. E. (1987). Incidental acquisition of word meaning from expositions with varied text features. *Reading Research Quarterly*, 22, 263–284.
- Horohov, J. E., & Oetting, J. B. (2004). Effects of input manipulations on the word learning abilities of children with and without specific language impairment. *Applied Psycholinguistics*, 25, 43–65.
- Horst, J. S., Parsons, K. L., & Bryan, N. M. (2011). Get the story straight: Contextual repetition promotes word learning from storybooks. *Frontiers in Psychology*, 2, 1–11.
- Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53, 983–997.
- Konopak, B. C. (1988a). Effects of inconsiderate vs. considerate text on secondary students' vocabulary learning. *Journal of Reading Behavior*, 20, 25–41.
- Konopak, B. C. (1988b). Eighth graders' vocabulary learning from inconsiderate and considerate text. *Reading Research and Instruction*, 27, 1–14.
- Konopak, B. C., Sheard, C., Longman, D., Lyman, B., Slaton, E., Atkinson, R., & Thames, D. (1987). Incidental versus intentional word learning from context. *Reading Psychology: An International Quarterly*, 8, 7–21.

- Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19, 304–330.
- Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24, 237–270.
- Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19–36). Hillsdale, NJ: Erlbaum.
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20, 233–253.
- Nash, M., & Donaldson, M. L. (2005). Word learning in children with vocabulary deficits. *Journal of Speech, Language, and Hearing Research*, 48, 439–458.
- Nation, K., Angell, P., & Castles, A. (2007). Orthographic learning via self-teaching in children learning to read English: Effects of exposure, durability, and context. *Journal of Experimental Child Psychology*, 96, 71–84.
- Oetting, J. B. (1999). Children with SLI use argument structure cues to learn verbs. *Journal of Speech, Language, and Hearing Research*, 42, 1261–1274.
- Oetting, J. B., Rice, M. L., & Swank, L. K. (1995). Quick incidental learning (QUIL) of words by school age children with and without SLI. *Journal of Speech and Hearing Research*, 38, 434–445.
- Pany, D., & Jenkins, J. R. (1978). Learning word meanings: A comparison of instructional procedures. *Learning Disability Quarterly*, 1, 21–32.
- Pany, D., Jenkins, J. R., & Schreck, J. (1982). Vocabulary instruction: Effects on word knowledge and reading comprehension. *Learning Disability Quarterly*, 5, 202–215.
- Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. Y. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research*, 37, 106–122.
- Rice, M. L., & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development*, 59, 420–429.
- SAS Institute. (2012). SAS/STAT (Version 12.1) [Computer software]. Cary, NC: Author.
- Schwanenflugel, P. J., Stahl, S. A., & McFalls, E. L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research*, 29, 531–553.
- Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals—Fourth Edition*. San Antonio, TX: The Psychological Corporation.
- Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 72, 95–129.
- Shu, H., Anderson, R. C., & Zhang, H. (1995). Incidental learning of word meanings while reading: A Chinese and American cross-cultural study. *Reading Research Quarterly*, 30, 76–95.
- Stahl, S. A., Burdige, J. L., Machuga, M. B., & Stecyk, S. (1992). The effects of semantic grouping on learning word meanings. *Reading Psychology*, 13, 19–35.
- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, 56, 72–110.
- Steele, S. C., & Watkins, R. V. (2010). Learning word meanings during reading by children with language learning disability and typically-developing peers. *Clinical Linguistics & Phonetics*, 24, 520–539.
- Swanborn, M. S. L., & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research*, 69, 261–285.
- Wagovich, S. A., & Newhoff, M. (2004). The single exposure: Partial word knowledge growth through reading. *American Journal of Speech-Language Pathology*, 13, 316–328.
- Wagovich, S. A., Pak, Y., & Miller, M. D. (2012). Orthographic word knowledge growth in school-age children. *American Journal of Speech-Language Pathology*, 21, 140–153.
- White, T. G., Power, M. A., & White, S. (1989). Morphological analysis: Implications for teaching and understanding vocabulary growth. *Reading Research Quarterly*, 24, 283–304.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

Appendix A

Rare Words Used in the Study

Nouns	Verbs
bombast	adduce
brogan	assuage
canard	aver
egress	bedaub
foible	convolve
onus	eschew
pathos	forfend
pelage	imbibe
precept	imbrue
runnel	impugn
surfeit	inveigh
tincture	niggle
torpor	refel
tyro	repine
unguent	truckle
welkin	wintle

Appendix B

Sample Multiple-Choice Items

Eschew

Syntactic Partial Word Knowledge Item

Is this word a(n)

- (a) person, place, thing, or idea
- (b) action word
- (c) don't know

General Semantic Partial Word Knowledge Item

The word eschew has something to do with

- (a) getting rid of or avoiding
 - (b) eating or drinking
 - (c) clothing or something worn
 - (d) travel
 - (e) don't know
-