Research Article

Exemplar Variability Facilitates Retention of Word Learning by Children With Specific Language Impairment

Jessica M. Aguilar,^a Elena Plante,^a and Michelle Sandoval^a

Purpose: Variability in the input plays an important role in language learning. The current study examined the role of object variability for new word learning by preschoolers with specific language impairment (SLI).

Method: Eighteen 4- and 5-year-old children with SLI were taught 8 new words in 3 short activities over the course of 3 sessions. Half of the children saw 3 identical objects corresponding to each new word during training (No Variability group); the other half of the children saw 3 different objects corresponding to each new word during training (High Variability group). Children completed vocabulary learning tests for objects seen during training and for new within-category objects that were never seen during training as a

test of category generalization. Learning was assessed the day after each training activity, and retention was assessed 3 weeks after the last training session.

Results: There were no group differences on trained or generalization items immediately following training sessions. However, children in the High Variability group demonstrated significantly better retention 3 weeks after experimental training.

Conclusion: These findings demonstrate that object variability facilitates retention of new word learning by children with SLI.

Supplemental Material: <https://doi.org/10.23641/asha.5583979>

Specific language impairment (SLI) is often associated with low vocabulary test scores relative to those seen in typical development (Gray, Plante, Vance, & Henrichsen, 1999; Hick, Joseph, Conti-Ramsden, Serratrice, & Faragher, 2002). This decrement is particularly problematic because vocabulary acquisition is widely acknowledged as a common component of children's linguistic abilities and academic achievement. Limited vocabulary and language ability have been associated with lower levels of social acceptance among peers in preschool (Gertner, Rice, & Hadley, 1994). Florit, Roch, and Levorato (2011) also found that preschool children's word knowledge accounted for differences in listening text comprehension. As children reach the school years, vocabulary plays an important role in academics. In particular, vocabulary is a significant predictor of word-reading ability (Kim, Al Otaiba, Puranik, Folsom, & Grulich, 2014; Oakhill & Cain, 2012; Suggate, Reese, Lenhard, & Schneider, 2014; Tunmer & Chapman, 2012). Vocabulary is also recognized as one of the five principal

components of writing, and the development of rich vocabulary is considered an essential step in becoming an effective writer (Baker, Gersten, & Graham, 2003). Clearly there is a critical role for vocabulary for linguistic abilities and academic achievement, making vocabulary development in SLI potentially consequential for educational achievement.

Several factors may contribute to relatively restricted lexicons for children with SLI. For example, these children require more exposures to learn a word, compared to their typically developing peers (Gray, 2004; Rice, Oetting, Marquis, Bode, & Pae, 1994). This slower rate of learning may itself limit the size of a child's lexicon for his or her age. This may cause even initially modest decrements in the lexicon to expand with age (Rice & Hoffman, 2015). It is also the case that children with SLI demonstrate reduced sensitivity to semantic features of words (Alt & Plante, 2006). As a result, even when words are present in the lexicon, they may be weakly encoded or weakly elaborated. This article addresses the encoding of lexical items along with their physical referents to determine whether this process can be facilitated during a brief treatment period for children with SLI.

Semantic Properties

An important component of learning includes creating and storing semantic representation of new words. This

^aThe University of Arizona, Tucson

Correspondence to Elena Plante: eplante@u.arizona.edu

Editor-in-Chief: Shelley Gray

Editor: Patricia Brooks

Received March 21, 2017

Revision received June 25, 2017

Accepted August 25, 2017

https://doi.org/10.1044/2017_LSHSS-17-0031

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

is the primary focus of this study. Sensitivity to semantic representations or “semantic features,” such as shape, color, material, and function, has been shown to influence word learning (Jones & Smith, 1998; Kemler Nelson, Russell, Duke, & Jones, 2000). Semantic representations play an important role in learning, because they give words meaning and allow learners to form categories. For example, to learn the nominal category “dog,” a child must come to understand that the word “dog” does not refer to just one dog, or type of dog, but to all the diverse examples of dogs that the child experiences.

Research suggests that children learning words and word categories are biased to attend to specific perceptual features of objects such as shape, material, or function. Several studies have provided evidence of a “shape bias” when children are asked to learn names of solid objects. However, children do not always attend to shape. For example, Jones and Smith (1998) found that children attended to both shape and texture when exemplar objects had eyes or shoes, indicating animacy. Other studies have also provided evidence that children use a “material bias” for non-solid objects (Perry, Samuelson, Malloy, & Schiffer, 2010). In addition, there is evidence that children generalize names for novel objects that share a similar functional property (Kemler Nelson et al., 2000). Therefore, it is clear that young children are capable of recognizing multiple aspects of object exemplars, and these aspects influence the likelihood that object exemplars will be seen as belonging to a single object class or to different classes. However, it is not the case that all properties of an object exemplar define the category. For example, a variety of sizes, shapes, colors, and textures are all permissible within the category of “dog.”

Children with SLI may have difficulty in differentiating between object attributes that define a category and those that are incidental. Several studies suggest that these children may have less developed knowledge of the visual properties for words they acquire. McGregor, Newman, Reilly, and Capone (2002) trained preschoolers on 20 object words and then tested the children on naming, drawing, and defining words. Results showed that children with SLI made more naming errors than their typically developing peers and that these errors were associated with less detailed drawings, suggesting weaker semantic representations. A similar finding was reported by Alt, Meyers, and Alt (2013) using a nonverbal task. These authors examined visual semantic knowledge of adults and children with normal and impaired language by asking participants to rate the “weirdness” for pictures of 20 different animals that varied in the typicality of their features (e.g., color, pattern, shape, and facial configuration). Results from this study showed that children with SLI gave significantly higher weirdness ratings to standard images (images that did not contain unusual visual features) than both children and adults with normal language. The authors suggested that these increased weirdness ratings might reflect weak visual semantic knowledge that restricts these children’s ability to judge normal category variants as acceptable exemplars.

This trait appears to have its roots in a weakness related to the early perceptual biases that facilitate word learning for typical children. Collisson, Grela, Spaulding, Rueckl, and Magnuson (2015) presented two forms of evidence that support this notion. In a similarity classification task, 3- and 4-year-old children with SLI were presented with an unlabeled target object and test objects that varied from the target on a single dimension (shape, color, or texture). When presented with a target object and asked which of several test objects went with it, children with SLI did not show an obvious shape bias on either task compared to their typical language peers. A similar result was found for a novel name extension task, in which children were presented with a target object that was labeled “this is a (novel name)” and instructed to “find another (novel name)” from the test items. The authors hypothesized that impairments in visual learning might partially underlie, or even compound, the linguistic weaknesses in SLI during word learning.

Weak recognition of important visual characteristics for object exemplars would hinder within-category semantic generalizations of word labels. Furthermore, if a child is impaired in simple visual learning of object classes, it should not be surprising that they are impaired in the multi-modal learning required to detect the relationship between objects’ shapes and their names (e.g., Alt et al., 2013; Alt & Plante, 2006; Alt, Plante, & Creusere, 2004; Gray, 2005; McGregor et al., 2002; Rice, Buhr, & Nemeth, 1990; Rice, Buhr, & Oetting, 1992; Rice et al., 1994). However, children with SLI can benefit from semantic cues during word learning. Gray (2005) used a range of semantic cues (e.g., superordinate, physical characteristic, action, part, and association) that were explicitly presented during word learning (e.g., “It’s green.”). This technique aided comprehension, suggesting an impact on semantic mapping between objects and labels. The weaknesses in visual semantic mapping suggest that treatment methods should directly address the object properties that define word–object mapping to improve word learning. Treatment methods that support this process should enhance children’s ability to generalize newly learned word labels to additional within-class members, beyond the specific exemplars trained.

Statistical Learning

Statistical learning is a general model of learning that applies to learning in both verbal and nonverbal domains. Statistical learning is the ability of individuals to extract statistical regularities from the world around them to learn about their environment. Statistical learning, particularly a sensitivity to conditional probabilities, is the suggested basis of learning many aspects of language, including recognition of lexical units and lexical-semantic mapping. Therefore, this type of learning is input-based but does not require explicit instruction. The advantage of this for children with impairments is that training methods grounded in statistical learning avoid metalinguistic-type instructions (e.g., it’s a dog because...), which may be beyond the

capabilities of young children. It also avoids explanations that may require use of vocabulary that is weak or unknown by these children (e.g., colors, shapes, relational terms).

Erickson and Thiessen (2015) suggest that different mental processes may account for the type of statistical learning needed to build the lexicon. For learning lexical labels, sensitivity to conditional probabilities aid in the identification of sounds and syllable co-occurrences that constitute individual words (i.e., identification of lexical units). Learners group adjacent sounds and syllables from speech to create a chunk or percept (e.g., a word). Once the percept is created, it is activated in memory. The level of activation decays over time, but it can be reactivated if the percept is encountered (e.g., another presentation of the same word). For example, in a training context, new words that are the target of training are typically presented multiple times. These words should show less decay than other words that also occur during training, but less frequently. Finally, interference also contributes to increasing the activation of consistent items relative to random items, because chunks that consistently share elements (e.g., training words) compete with other syllable strings that might make up sentences. As a result of this process, consistently co-occurring sound sequences that reflect recurring lexical items will be learned over less frequent or random groupings.

Erickson and Thiessen (2015) further proposed a second process, extraction, that may be critical to word learning. Extraction refers to an individual's ability to hold distinct elements in working memory and bind them together into a single chunk. For vocabulary, this would involve binding lexical labels with the semantic properties that define the word. This process is sensitive to conditional statistical information in the input (e.g., the co-occurrence of a word and the collection of particular physical features that define that word), and it is thought to rely on both attention and working memory.

Variability in the linguistic input is known to facilitate multiple forms of statistical learning (Gómez, 2002; Plante et al., 2014; Richtsmeier, Gerken, Goffman, & Hogan, 2009; Torkildsen, Dailey, Aguilar, Gómez, & Plante, 2013). Several experimental studies of statistical learning have provided strong evidence that learners notice patterns in their perceptual input (e.g., Aguilar & Plante, 2014; Gómez, 2002; Gómez & Maye, 2005; Grunow, Spaulding, Gómez, & Plante, 2006). Using these types of learning paradigms, researchers have identified several "principles" that can aid learning in therapeutic contexts (e.g., Alt, Meyers, & Ancharski, 2012). One principle of learning that is most important to this study concerns exemplar variability. The idea behind this principle is that, if a clinician wants to teach a target form, the clinician should vary everything in the input to the child except for the target structure. Variation around the target is thought to make the target of learning perceptually salient and therefore easier to learn (Gómez, 2002; Torkildsen et al., 2013).

Gómez (2002) first examined the effects of variability in the input for learning nonadjacent dependencies of an

artificial grammar, similar to English grammar (e.g., aXb, where a predicts b, but the intervening X element is free to vary as in is playing, is jumping, is running). Both adults and infants who heard 24 exemplars of the X element during exposure to the aXb-type grammar learned significantly better than those who heard fewer X elements. Moreover, the High Variability groups demonstrated generalization of learning to untrained exemplars.

Many artificial grammar learning studies have demonstrated significant effects of learning and generalization of grammars by both children and adults after only minutes of exposure (Gómez, 2002; Gómez & Maye 2005; Grunow et al., 2006; Torkildsen et al., 2013). Torkildsen et al. (2013) examined variability and grammar learning by adults with normal language and language impairment. Results from this study indicated similar patterns of learning for participants, normal or impaired, who received training in the high-variability condition. Plante et al. (2014) integrated the variability principle into a treatment context. They used a conversational recast treatment to teach grammatical morphology to preschoolers with SLI. During treatment, clinicians recasted children's morphological target structure with 24 or 12 different sentence contexts. Children in the High Variability group (i.e., 24 exemplars) learned better than children in the No Variability group (i.e., 12 exemplars, each presented twice).

These studies all report evidence of learning for both typical and impaired groups under high-variability input conditions. In each study, however, the language-impaired group showed a slightly attenuated pattern of learning relative to the normal language group. Therefore, it appears that high-variability input facilitates learning but does not make it completely normal (Aguilar & Plante, 2014; Plante, Bahl, Vance, & Gerken, 2011; Torkildsen et al., 2013). Nonetheless, individuals with language impairment clearly demonstrate improved sensitivity to statistical patterns in the input when lexical variability is high.

Although work to date with language impairment has targeted morphological learning, there is evidence that input variability can facilitate aspects of word learning, at least for typically developing children. Perry et al. (2010) taught 18-month-olds names of new categories (e.g., buckets) with three very similar objects in the Low Variability group and three dissimilar objects in the High Variability group in nine weekly visits. Children in the High Variability group not only learned more of the trained object categories but also generalized learning to new unfamiliar words. Importantly, they demonstrated continued vocabulary growth for 1 month after training, suggesting that their experimental experience had improved their overall capacity for word learning. The authors suggested that object variability helped these 18-month-olds use learning biases (e.g., the tendency to encode shape, material, function) to map semantic categories in ways that promoted learning of new words.

Other evidence appears to run counter to this finding. Twomey, Ranson, and Horst (2014) examined the effects of encountering variable exemplars on 2.5-year-old children's gradual, long-term word learning. In their initial

experiment, children learned nonword labels when the same object exemplar was presented repeatedly or when multiple object exemplars that differed in color were presented. Then, after a 5-min delay, children were tested to assess their retention of learning. Results showed that all children did well during training, but only the children who encountered variable exemplars retained learning after the short delay. This is consistent with the idea that variable input assists learning, at least in terms of retaining learning over time. Experiment 2 replicated and extended these findings by exploring the effect of greater within-category variability on children's retention of word learning. In this experiment, 2.5-year-old children were taught new words with variable objects that differed in color only (narrow condition) or objects that differed in color, size, and texture (broad condition). Again, results indicated that children in both groups did well on the initial test, but only children in the narrow condition retained learning after the delay. Although results for the narrow condition are consistent with the variability principle, those from the broad condition were not. The researchers suggested that too much variability might have added greater attentional demands to the task for the young children. When the objects were more variable, fewer resources may have been available for committing the name-object associations to long-term memory. If so, as children mature and their cognitive capacity expands, higher levels of input variability may no longer prove challenging. Input variability may instead serve to differentiate between object features that are extraneous to class membership versus those that are central, facilitating learning in somewhat older children.

Word learning by young children has also been examined using contextual variability. Vlach and Sandhofer (2011) examined how changes in contextual information across learning and testing affected 2.5-, 3-, and 4-year-old children's word learning and generalization. In this study, all children participated in a novel noun generalization task in which background context was manipulated. During the learning phase of training, children saw object exemplars against a single colored background or with multiple backgrounds of different colors. The object exemplars used to train each word were three novel objects that differed in color, texture, and perceptual features but had the same shape. During each presentation, the object was presented against either the consistent or variable backgrounds and labeled two or three times. Results showed that both the 3- and 4-year-olds learned in both the one background context condition and the multiple background context condition. However, the 2.5-year-old children only learned in the one context condition and did not benefit from context variability, suggesting a developmental effect for input variability. These results are similar to those of Twomey et al. (2014), who found that too much variability did not aid learning at this young age. Because the objects used in training for the Vlach and Sandhofer (2011) study already varied in color, texture, and perceptual features, perhaps varying the context as well was too much variability for these young children. However, by the age of 3 years, children

begin to show benefits of both object and context variability that facilitate word learning (i.e., generalization).

Treatment of Vocabulary Deficits in SLI

It is clear that variability plays an increasingly important role in helping young learners recognize the patterns in their input that are critical for learning words. Although variability and word learning have not been studied with individuals with SLI, it is possible that incorporating object variability into language intervention may improve the efficacy and effectiveness of treatment.

The purpose of this study is to test the variability principle derived from statistical learning studies in a word-learning treatment for preschool-age children with SLI. This study will use methods similar to Perry et al. (2010) in terms of its focus on the variability of the object exemplars used to represent nouns to be learned. That is, a training condition that uses object exemplars that are identical in appearance is contrasted with a condition in which physically dissimilar objects are used. The hypothesis is that children with SLI in a high exemplar variability condition will learn more words presented within the experiment than children in a no exemplar variability group. Learning will be defined as generalization of the object label to new object exemplars of the same class. Variable exemplars should lead children to focus on the features common to each object exemplar associated with a single label. This focus on common features should allow children to form a category that includes many different object exemplars rather than producing a one-to-one mapping of words to a particular object form. This, in turn, should allow children to generalize to new within-category exemplars of each word learned. This focus on generalization items, rather than trained items, is consistent with the idea that true learning in treatment is reflected by performance on untrained rather than trained exemplars (Kamhi, 2014).

Method

Participants

Eighteen children (14 boys, four girls) with SLI participated in the experiment. Nine were assigned to a no-variability treatment condition, and nine were assigned to a high-variability treatment condition. Children ranged in age from 4;4 to 5;9 (years;months). All children were reported by their parents to be native speakers of English. Two children also heard Spanish spoken at home. One child, originally in the high-variability condition, withdrew from the study after the training phase and did not complete the final experimental posttests. He was replaced with another boy who completed all study phases. All parents provided informed consent for their child's participation.

Sixteen of the 18 children received speech-language therapy to treat grammatical morphology during the time of their participation in this study. Fourteen of those 16 children also received therapy for speech sound disorders concurrently. Two children received therapy for

speech sound disorders only but also evidenced language deficits on standardized testing (described below). All children participated in preschool classroom activities when not receiving therapy. The preschool program consisted of a general language enrichment program with weekly themes. Each theme included a set of related vocabulary words. None of those words were trained in the vocabulary treatment.

Prior to the start of the experiment, children were administered a battery of tests to confirm their diagnosis of language impairment and to describe their language and cognitive skills (see Table 1). All children had nonverbal cognitive skills that fell above the range for intellectual disability (criterion of 70, +1 *SEM*, or 75) on the nonverbal scale of the Kaufman Assessment Battery for Children–Second Edition (Kaufman & Kaufman, 2004). This criterion is consistent with the classic definition of SLI that rules out other handicapping conditions (in this case, intellectual disability) as a cause of poor language skills (see Plante, 1998, for a discussion of issues surrounding IQ in SLI). In addition, all children scored below an empirically derived cutoff score of 87 (Greenslade, Plante, & Vance, 2009), indicative of language impairment on the Structured Photographic Expressive Language Test–Preschool 2 (Dawson et al., 2005). Children also received the Goldman-Fristoe Test of Articulation–Second Edition (GFTA-2; Goldman & Fristoe, 2000) to describe their articulation skills. There were no significant differences between the High Variability and No

Variability groups for any of the test scores (two-tailed *t* test, $p > .5$ in all cases; Kaufman Assessment Battery for Children–Second Edition, $d = 0.43$; Structured Photographic Expressive Language Test–Preschool 2, $d = 0.07$; GFTA-2, $d = 0.08$). All children also passed a pure-tone hearing screening at 1000, 2000, and 4000 Hz at 25 dB HL.

Children were also administered the Peabody Picture Vocabulary Test–Fourth Edition (PPVT-4; Dunn & Dunn, 1997) and the Expressive Vocabulary Test–Second Edition (Williams, 2007) to describe their vocabulary skills prior to training. Previous research has shown significant relationships between initial vocabulary knowledge and word learning (Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009). Therefore, receptive vocabulary scores were used to assign children to groups by ranking their scores on the PPVT-4 and alternating assignment to each treatment condition. This was done to assure roughly equal vocabulary scores across groups. Consequently, there were no significant differences between the High Variability and No Variability groups for the vocabulary test scores (two-tailed *t* test, $p > .5$; PPVT-4, $d = 0.37$; Expressive Vocabulary Test–Second Edition, $d = 0.15$).

Materials

Target Words

Potential target words were 20 one- and two-syllable nouns that were anticipated to be unfamiliar to the children

Table 1. Demographic and test data for children in high- and no-variability conditions.

Participant	Age (months)	Gender	Mother's education (years)	K-ABC2	SPELT-P2	PPVT-4	EVT-2	GFTA-2
<i>High Variability group</i>								
H-1	57	F	14	98	56	80	74	108
H-2	52	F	12	81	74	91	97	102
H-3	67	M	16	82	56	97	82	60
H-4	57	M	16	91	58	94	87	70
H-5	58	M	13	87	53	77	75	74
H-6	58	M	17	109	68	122	109	83
H-7	54	M	14	95	65	88	90	75
H-8	63	M	15	109	53	94	96	50
H-9	69	M	12	99	67	109	97	79
<i>M</i>	59.4		14.3	94.5	61.1	94.7	89.7	77.9
<i>SD</i>	5.7		1.8	10.4	7.6	13.9	11.4	18.4
<i>No Variability group</i>								
N-1	54	F	14	120	63	88	88	73
N-2	52	F	14	93	56	94	97	93
N-3	63	M	13	98	77	93	99	73
N-4	55	M	17	111	76	85	96	89
N-5	53	M	11	100	80	85	86	97
N-6	58	M	14	96	79	98	91	80
N-7	60	M	13	91	58	73	85	81
N-8	56	M	15	116	75	112	93	78
N-9	64	M	13	77	53	78	81	51
<i>M</i>	57.2		13.7	100.2	68.5	89.5	90.6	79.4
<i>SD</i>	4.3		1.6	13.4	10.8	11.5	6.1	13.6

Note. All test scores are standard scores with a mean of 100 ($SD = 15$). K-ABC2 = Kaufman Assessment Battery for Children–Second Edition (nonverbal scale); SPELT-P2 = Structured Photographic Expressive Language Test–Preschool 2; PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition; EVT-2 = Expressive Vocabulary Test–Second Edition; GFTA-2 = Goldman-Fristoe Test of Articulation–Second Edition; F = female; M = male.

in this study. Of those 20 words, eight were chosen as target words for training for each child. These words referred to physical objects, which are described below. The trained words and their sublexical characteristics are reported in Table 2. We tested for significant group differences in the number of syllables ($\chi^2 = 0.25, p = .6201$), word frequency ($t = 0.03, p = .98$), neighborhood density ($t = 0.97, p = .16$), and phonotactic frequency ($t = 1.07, p = .57$) for words selected for training. None of these differed significantly between High Variability and No Variability groups.

Objects

Target objects were selected from a variety of craft and hardware items. Examples are provided in Figure 1. We selected 20 objects whose labels were likely to be unfamiliar to preschool children. Objects varied in terms of their physical attributes (e.g., size, color, material, shape). Children in the No Variability group were exposed to three identical object exemplars, which were judged to be most prototypical by two adults, because semantic judgments are better for prototypical objects (Alt et al., 2013). The children in the High Variability group saw the same prototypical object exemplar that was seen in the no-variability condition, and they also saw two additional object exemplars that were visually very different from each other.

Vocabulary Selection Tests

Several tests were developed for use in this study. The format of all experimental tests was similar to the PPVT-4 in that a test booklet was used and pages contained four pictures of potential training words. The tester said the name of one of these words, and the children pointed to the picture of the corresponding object. All tests were scored as the number of correctly selected pictures corresponding to named objects. The first was a vocabulary selection test containing all possible target words to be used in the experiment. This was used to select target words for experimental training and test children's learning of trained words after each activity. The vocabulary selection test used pictures of the prototypical object exemplars

Figure 1. Example of no-variability and high-variability objects used during training and generalization testing. Note that the high-variability hinges were of different sizes, shapes, and metal finishes. The high-variability tassels were of different sizes, colors, and textures. All children were tested on generalization items that differed from those seen during training.

	Training			Testing
No Variability Hinges
High Variability Hinges
No Variability Tassels
High Variability Tassels

to identify words that children could not recognize. Children heard 20 potential target words, one at a time, and were asked to point to the picture corresponding to that word in a four-picture array. Eight words unfamiliar to the child (i.e., words that children did not correctly identify) were selected for training.

Learning Tests

Each child completed individualized word identification tests the day after each vocabulary activity. The tests were designed to measure acquisition and retention of each child's target words 1 day after training. This examined learning of exemplars that were seen in both high- and no-variability conditions and generalization to new exemplars that were never seen in training. Generalization items were within-category items that the children never saw during training (see Figure 1 for examples). Each test booklet contained 16 test items corresponding to the eight items seen

Table 2. Words selected for training and their sublexical characteristics.

Words	Word frequency ^a	Sublexical characteristics		Number of children trained on each word	
		Biphone average	Neighborhood density	High-variability condition	No-variability condition
Doily	3.135	0.2005	0	7	8
Burlap	5.081	0.2006	0	5	4
Mesh	7.578	0.3344	6	1	1
Twine	5.781	0.2509	4	4	5
Tote	6.875	0.3355	30	7	8
Sequin	3.761	0.1682	0	9	8
Bobbin	5.198	0.2020	2	7	7
Bolt	8.895	0.2529	14	7	8
Tassel	3.783	0.2025	6	9	9
Hinge	6.862	0.2540	4	8	8
Streamer	6.664	0.1719	3	8	6

^aLog frequency word estimates from the Hyperspace Analog to Language database (Balota et al., 2007).

during training. Eight items measured learning of trained exemplars (i.e., exemplar seen in both high- and no-variability conditions). The foils for these eight items consisted of photographs of other objects seen during the training tasks so that both target words and foils were visually familiar to the child. Thus, there was no difference in children's training experience with words that served as correct and foil items, which can inflate word identification test scores (cf. Axelsson & Horst, 2013). Eight items measured learning of untrained exemplars (i.e., generalization items). These included new exemplars of the target item, and the foils included untrained exemplars of other trained words. These were customized for each child so that each child was tested only on those vocabulary words included in his or her own treatment.

Procedure

The experiment took place in 5 days spaced over a 6-week period. The sequence of events is illustrated in Figure 2. The vocabulary selection test was used to identify words unknown to each child that could be trained during the experiment. Training occurred over 3 days within a 3-week period. The first two training sessions occurred consecutively in the first week, and the third happened during either Week 2 or Week 3. Children's learning was tested the day immediately following each of the three training sessions. Testing always preceded additional training so that test score reflected retention from the prior training session, rather than immediate retention of the same day's learning. A final test session was used to determine retention of vocabulary words after training was administered at Week 6. The individuals testing the children did not participate in training, maintaining experimental blinding for both testers and training clinicians.

Testing

The vocabulary pretests were administered to all children on the first day of Week 1 to select target words. Children also completed individualized testing of the specific words for which they received training on the day after each training session and again approximately 3 weeks

($M = 20.8$ days; range 19–22 days) after training to measure retention.

Learning Activity

Previous word-learning studies have reported anywhere from one to 240 presentations of each target word to children with SLI per training session (Alt & Plante, 2006; Gray, 2005; Gray, Pittman, & Weinhold, 2014; McGregor et al., 2002; Rice et al., 1992). Studies utilizing exemplar variability with typical children have not reported total number of target words presentations per session (Perry et al., 2010). In this study, clinicians presented each target word six times in each activity for a total of 18 presentations of each word in three sessions over 3 weeks. Previous studies have shown that children with SLI can begin the process of semantic mapping with as few as three to 10 presentations (Alt & Plante, 2006; Rice et al., 1994). Eighteen presentations of each of the eight training words over three sessions were intended to move the children past the initial stage of learning, without creating a ceiling effect for both groups.

Children completed three training activities over a 3-week period. The average length of each activity was 12 min 38 s, with a range from 7 min 38 s to 20 min 24 s. Children completed the first two activities on two different consecutive days ($M = 1$ day, $SD = 0$) during Weeks 1 and 2 of the 6-week program. A preliminary examination of the data after the first two training sessions indicated that most children might benefit further from a third day of training without reaching ceiling performance. This was added during the third week of the study. The time between the second and third training session ranged from 7 to 12 days ($M = 8.8$ days, $SD = 2.17$). The order of the activity presentation for these first 2 days was counterbalanced across conditions.

Activities included building a child-sized robot and going on a pirate treasure hunt. These activities were designed to be engaging for the children, gender neutral, and allowed for the target words to be presented in a naturalistic and meaningful context. For the third session, children repeated either the robot activity or the pirate treasure hunt, with the activities counterbalanced across groups. Four children in the no-variability condition and four children in the high-variability condition repeated the robot

Figure 2. The sequence of experimental methods. Note that Days 1 and 2 are consecutive, as are Days 3 and 4. However, between 2 and 3 weeks passed between Days 2 and 3. Retention testing occurred 6 weeks from the first training day.

activity, and five children in the no-variability condition and five children in the high-variability condition repeated the pirate treasure hunt activity. All training sessions occurred during Weeks 1–3, and generalization test sessions occurred during Weeks 1–3 and 6 of the 6-week program. Each clinician provided training with both activities, and different clinicians trained different children over the 3 days. This assignment was based on the availability of the clinician and the child.

Clinicians began training by looking at the printed instructions for the day's activity (i.e., treasure hunt clues, robot-building instructions) with the child. Review of the instructions to the child provided the first presentation of each target word by reading the instruction sentences to the child. The clinician presented the three corresponding object exemplars for each word immediately before, during, or immediately after the first presentation of the target word. Both the clinician and children were allowed to handle all three corresponding exemplars. However, children only used one of the three objects to complete each activity, and the other two were put away before the presentation of the next word and object exemplars.

The first four presentations included the following: a sentence from the activity directions, a statement (e.g., "Here's a _____.", "This is a _____.", "This is my favorite _____."), an imperative command (e.g., "Get a _____.", "Find the _____."), and a question (e.g., "Which ____ do you want?", "Which ____ should we use?", "Which is your favorite ____?"). This was followed by a prompt for the child to produce the word ("What is this called?"). After attempting the target word, the clinician provided feedback for a fifth presentation of the target word (e.g., "Yes, this is a _____.", "No, this is a _____."). The order of object and word presentations was not fixed for either activity, and the clinician was allowed to present them in any order. This allowed for clinicians to follow the child's interest in the objects in a more naturalistic way than a fixed order would have allowed. At the end of each activity, the child received the sixth presentation of each word while taking exemplars off the robot in the robot activity and putting the exemplars in the treasure chest in the pirate activity.

Clinicians randomly presented target words in each activity for all children and tracked the order of presentation by filling in the blanks on activity instructions. Clinicians used a paper record form to track the number of times each word was presented, the linguistic context of each presentation (i.e., statement, question, imperative), and the child's attempts to use target words, either elicited or spontaneous.

In order to facilitate replication of treatment methods, as well as comparisons across studies, Warren, Fey, and Yoder (2007) recommended standard reporting of a number of critical treatment parameters. In their framework, "dose" refers to the events thought to affect change in treatment. As described above, there were two types of potential treatment doses in this study. The most frequent dose consisted of the clinician's presentation of a word and its

referent (a receptive dose for the child) presented to the child in the context of an ongoing activity. In addition, each child was asked to attempt to name each object once per session (an expressive dose for the child). This expressive practice may also have served as a dose because previous research has provided evidence that having children generate words during learning produces stronger retention of words (Vlach & Sandhofer, 2012). However, children's attempts at using each word were not always correct when attempted, potentially undermining the effectiveness of word production as a dose. Therefore, this attempt was not counted as a true dose. Dose number refers to the number of properly administered teaching episodes during a single intervention session (Warren et al., 2007).

In this study, there were six receptive doses per word trained (i.e., target words spoken by clinician) per session. Dose frequency is the number of times treatment is provided per day and per week. In this study, dose frequency was once per day. Total intervention duration, which is the time period over which intervention is presented, was 3 weeks. The cumulative intervention intensity is calculated as $\text{Dose} \times \text{Dose Frequency} \times \text{Total Intervention Duration}$. In this study, this was relatively low at 18 teaching episodes per word.

Administration Fidelity and Reliability

Clinician Training

Four research clinicians provided the treatment. Two were certified speech-language pathologists, and two were novices to the field of speech-language pathology (i.e., individuals with no formal training in the field). Clinicians received training on how to present target words/exemplars for each activity and track data using the record form. During training, clinicians had the opportunity to practice presenting target words and exemplars within each activity. All clinicians provided training in both activities with different children over the 3 days. This assignment was random and based on availability of the clinician and the child. All testing was completed by two additional adults who were blind to the treatment. One was a certified speech-language pathologist, and the other was an undergraduate research assistant.

Procedural Fidelity

Procedural fidelity reflects the degree to which training was administered as intended and was coded as point-to-point reliability for 15 sessions (29% of all sessions). Two individuals who did not participate in training the child scored procedural fidelity. These sessions were distributed randomly across the experimental training sessions and represented all child participants, both activities (i.e., robot and pirate treasure hunt), and each of the three training sessions. Coding of procedural fidelity was calculated as total presentations of target words (i.e., 48 presentations, six presentations per eight target words) and total number of prompts to produce the target word (i.e., eight prompts, one prompt per target word). Omissions and additions of

presentations of the target word were deducted from the base of 48 presentations in each activity. Omissions of prompts were deducted from the base of eight prompts in each activity. Procedural fidelity for number of presentations was 95% (range = 92%–100%) and 93% (range = 87%–100%) for number of prompts.

One child in the high-variability condition missed seeing two of the three exemplars for one of his eight target words. The objects had inadvertently been removed from the pirate treasure hunt area prior to the child encountering them. The object exemplars that he did see were the exemplar that was trained in both high- and no-variability conditions. He saw all the other physical referents in that session and the correct number of physical referents in the other two sessions. He also received all doses for all three sessions. For the generalization test following this activity with the missing object, the child correctly identified the trained exemplar, indicating there was no adverse effect of this single omission. Therefore, this child was included despite this one omission in his treatment.

Scoring Reliability

A live double scorer coded reliability for 16 of the generalization test sessions following training (Tests 1, 2, and 3). Point-to-point reliability was calculated for agreement concerning child's correct identification of the target item. Scoring reliability was 97% (range 87%–100%).

Results

The hypothesis was that children in the High Variability group would learn more words than children in the No Variability group, as evidenced by their ability to correctly identify new exemplars of the word referents. Table 3 presents the means and standard deviations for the accuracy of the eight generalization items for all subjects on each test. The primary measure of learning is reflected by accuracy on generalization items rather than training items. Given that training items may reflect familiarity due to exposure during training rather than learning per se, true

learning is reflected by performance on untrained exemplars (Kamhi, 2014). The response to treatment by each participant is provided in Supplemental Material S1.

Generalization items were analyzed using a mixed analysis of variance with group (High Variability, No Variability) as the between-groups factor and test day (Test 1, Test 2, Test 3, retention test) as the within-group factor. The main effect for group was not significant, $F(1,16) = 0.84$, $p = .37$, $\eta^2 = .05$. There was a significant main effect for test day, $F(3,48) = 8.23$, $p = .0001$, $\eta_p^2 = .35$, and a significant Group $\times$ Test Day interaction, $F(3,48) = 4.49$, $p = .006$, $\eta_p^2 = .22$.

These significant effects would be considered small to medium effects for psychological research (Cohen, 1988). Calibrating effect sizes for clinical research is more subjective, requiring the clinician to balance the size of the effect with the cost of obtaining it and the benefit to the client. In terms of raw size, we would consider this a modest effect. However, the absolute size of the effect is offset by the very short intervention period. Achieving a significant effect in 3 days indicates that the cost of the improvement is relatively low. Moreover, the benefits are likely to accumulate if vocabulary training is extended.

The analysis of variance was followed by a Fisher's least significant difference post hoc, which revealed that the High Variability group correctly identified more generalization items than the No Variability group at retention test ($p < .006$), but not during training (see Figure 3).

A least squares analysis indicated that the High Variability group correctly identified more generalization items at Test 3 compared to Tests 1 ($p = .002$) and 2 ($p < .002$, two tailed), but Tests 1 and 2 did not differ. In addition, the High Variability group correctly identified more generalization items at the retention test compared to Tests 1 ($p = .0001$) and 2 ($p < .0006$), but not Test 3. The No Variability group only showed an increase in correct >identification of generalization items from Test 1 to Test 2 ($p = .05$) and to Test 3 ($p = .02$), but no additional gains occurred from Tests 2 to 3. Retention testing was not different from Test 1, Test 2, or Test 3.

Table 3. Means and standard deviations of generalization items for the participant groups.

Test	Group			
	High Variability		No Variability	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	3.00	1.12	3.11	1.36
2	3.67	1.73	4.00	1.00
3	4.78	1.56	4.33	1.22
Retention test	5.33*	1.58	3.44*	1.33

Note. There are a total of eight generalization items. Retention testing occurred on average 21.8 days ($SD = 1.05$) after last training session.

* $p < .05$.

Figure 3. Words learned at each test for High Variability and No Variability Groups. *Indicates a significant difference between groups.

We examined the relation between factors that could impact word learning (pretraining vocabulary skills, mother's education as a proxy for socioeconomic status). There were no significant correlations between receptive or expressive vocabulary nor mother's education and retention of vocabulary after training. Likewise, there was no relation between the average neighborhood density or phonotactic frequency of the words targeted for learning and the number of words that were generalized at retention.

Discussion

We predicted that high exemplar variability would produce better learning of targeted lexical items in a therapeutic context than no exemplar variability. The main finding of this study broadly supports this hypothesis for preschool children with SLI. Children in the High Variability group retained more words 3 weeks after training than did children in the no-variability condition. Better learning of novel vocabulary under high exemplar variability training conditions is consistent with Perry et al. (2010), who provided the first demonstration of the benefits of object exemplar variability for lexical learning. These authors showed that as few as three variable exemplars could facilitate word learning. As in that study, the current study also used three variable exemplars in the high-variability condition to teach children with SLI new words. The children in the High Variability group showed significantly stronger performance than the No Variability group on new generalization items at retention testing. These results suggest that the statistical learning principle of exemplar variability applies to the domain of lexical learning and that including exemplar variability as a treatment component has the potential for improving the efficacy of lexical treatment for children with SLI.

The current findings complement prior work by others for typically developing children (Perry et al., 2010; Twomey et al., 2014) and extend them to a clinical population characterized by language deficits. However, it is important to note the differences between studies, particularly that of Perry et al. (2010), which was the most similar to the current study. That study examined the role of exemplar variability for word learning by typically developing toddlers across a longer period of training. In addition, Perry et al. reported learning scores that combined the test scores from Weeks 3, 6, and 9. Therefore, it is unclear whether children who experienced object variability learned better than the children who did not on all or some of the individual learning tests in the Perry et al. study. In this study, both groups performed similarly on the tests given during training. Therefore, the effects of exemplar variability were not apparent during experimental training but emerged weeks after training on retention testing. This effect of retention of learned words using high exemplar variability provides important considerations for clinicians providing lexical interventions. Although the positive effect of high-variability input on

learning may not be obvious during training, it has the ability to produce lasting effects of learning that is retained.

The statistical learning framework predicts that learners will track the regularly co-occurring patterns in the input. Variability for elements extraneous to these recurring patterns serves to make these patterns salient (e.g., Gómez, 2002; Plante et al., 2014; Torkildsen et al., 2013). We have shown here that variability of the object representations appears to highlight the common category features of item classes apparent to children, facilitating generalization of word labels to other in-class items. Thus, the variability principle, already known to facilitate grammatical learning in language impairment (Grunow et al., 2006; Plante et al., 2014; Torkildsen et al., 2013), is extended here to the domain of vocabulary.

In this study, children in the High Variability group demonstrated retention of learning new words with only six presentations of each target word in each of three relatively short activities. Thus, learning was achieved efficiently, without prolonged training. This efficiency is consistent with the general nature of statistical learning, at least in typical learners, in that measurable learning occurs on the time scale of minutes rather than weeks or months. Our children with SLI likewise demonstrated measurable learning after a single training session, having acquired an average of three new words after 18 exposures to each. Previous studies targeting word learning for children with SLI typically report using only one object exemplar per word with considerably more presentations of each target word across several training sessions or training blocks (Alt & Plante, 2006; Gray & Brinkley, 2011; Gray et al., 2014; Hoover, Storkel, & Hogan, 2010; Suanda, Mugwanya, & Namy, 2014). On the high end of dose number, Gray et al. (2014) trained children with SLI and typically developing children matched on age or vocabulary skill level. They used 16 nonwords with only one physical referent per word. Children with SLI required approximately 40 trials per word to learn approximately half the words trained (a level equivalent to learning in this study). Similar levels of learning in our high-variability training make this more efficient than high-repetition training using a single physical referent.

When asked to map objects to labels, children are biased to attend to specific perceptual features of objects such as shape or material when learning novel names. For example, several studies have provided evidence of a "shape bias" when children are asked to learn names of solid objects (Deák, Ray, & Pick, 2002; Graham, Williams, & Huber, 1999; Jones & Smith, 2002). Under some circumstances, children can attend to multiple perceptual features simultaneously (Alt et al., 2004; Alt & Plante, 2006; Jones & Smith, 2002). In the current study, the perceptual features of the objects in general were not overtly highlighted (e.g., plastic as the material for mesh or metal for hinge). Still, children in the High Variability group extended learning of words to new within-class items, suggesting that they had sufficient knowledge of semantic representations to drive the formation of categories. Others (Twomey et al., 2014; Vlach & Sandhofer, 2011) have shown that typically developing children benefit

from input variability by the age of 3 years, although results are mixed at younger ages (Perry et al., 2010; Twomey et al., 2014; Vlach & Sandhofer, 2011). In this study, our children with SLI appeared to operate more like typically developing children than younger typical children in that the children with SLI also showed benefit from object variability.

This study included children whose mean single-word vocabulary test scores were comparable to other recent samples of English-speaking preschool children with SLI (Fey, Leonard, Bredin-Oja, & Deevy, 2017; Haebig, Sterling, & Hoover, 2016; Meyers-Denman & Plante, 2016; Oetting, McDonald, Seidel, & Hegarty, 2016; Snowling, Duff, Nash, & Hulme, 2016; Vuolo, Goffman, & Zelaznik, 2017; Wray, Norbury, & Alcock, 2016). This suggests likely generalization to the broader SLI population. Whether similar effect sizes would occur when vocabulary is more severely impaired is open to speculation. Our results indicate no correlation between single-word vocabulary measures and the degree of learning, which would argue that the benefit of high-variability input over low-variability input would still benefit children with more severe vocabulary deficits. In this case, we might expect children with particularly low vocabulary levels to perform more poorly overall but still derive benefit from variable input. Another possibility is that the relative boost from high-variability input may fade as vocabulary test scores decrease into the severe range. Vocabulary scores covary with nonverbal IQ scores (Rice & Hoffman, 2015), and nonverbal IQ is a mediator for vocabulary learning (Kan & Windsor, 2010). It may be that when single-word test scores are very low, variability can be less facilitative, not because of the effects of a restricted lexicon per se, but because of cognitive limitations. This is consistent with the findings that very young children, who presumably have less developed cognitive resources than older children, do not benefit as consistently from input variability (Twomey et al., 2014; Vlach & Sandhofer, 2011).

Conclusion

The current study provides important insights into the nature of word-learning interventions for children with SLI. First, children in both groups demonstrated learning that generalized in the short term to untrained exemplars. Generalization to untrained exemplars is the gold standard for true language learning. In contrast, familiarity with items seen during training may reflect training effects rather than true generative use (Kamhi, 2014). Children in the high-variability condition outperformed children in the no-variability condition at retention. The findings suggest that teaching children new lexical items using high-variability exemplars may not affect learning in the short term, but that the benefit emerges with time. This is particularly important because it confirms Kamhi's (2014) hypothesis that if clinicians only measure performance on trained items, they may be missing critical aspects of treatment efficacy. Second, children were able to learn new words that reflected a range of phonological characteristics known to facilitate

word learning (e.g., frequency, neighborhood density). This finding is noteworthy because the procedures used here are easier to implement in real-world settings than limiting training to those words in which frequency and neighborhood density characteristics facilitate learning. Finally, teaching children new words using high-variability exemplars in interactive learning contexts supports children's interest, attention, and conversational interactions. Therefore, this study provides positive results for a word-learning intervention that can be easily implemented by clinicians.

Acknowledgment

This work was funded by National Institute on Deafness and Other Communication Disorders Grant 1R21DC014203 awarded to the University of Arizona (Elena Plante, principle investigator). The work was completed in partial fulfillment of the doctoral degree for the first author and was previously presented in 2016 at the Symposium for Research in Child Language Disorders.

References

- Aguilar, J. M., & Plante, E. (2014). Learning of grammar-like visual sequences by adults with and without language-learning disabilities. *Journal of Speech, Language, and Hearing Research*, 57, 1394–1404.
- Alt, M., Meyers, C., & Alt, P. M. (2013). Using ratings to gain insight into conceptual development. *Journal of Speech, Language, and Hearing Research*, 56(5), 1650–1661.
- Alt, M., Meyers, C., & Ancharski, A. (2012). Using principles of learning to inform language therapy design for children with specific language impairment. *International Journal of Language & Communication Disorders*, 47, 487–498.
- Alt, M., & Plante, E. (2006). Factors that influence lexical and semantic fast mapping of young children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 49, 941–954.
- 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(2), 407–420.
- Axelsson, E. L., & Horst, J. S. (2013). Testing a word is not a test of word learning. *Acta Psychologica*, 144, 264–268.
- Baker, S., Gersten, R., & Graham, S. (2003). Teaching expressive writing to students with learning disabilities research-based applications and examples. *Journal of Learning Disabilities*, 36, 109–123.
- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, 39, 445–459.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Collis, B. A., Grela, B., Spaulding, T., Rueckl, J. G., & Magnuson, J. S. (2015). Individual differences in the shape bias in preschool children with specific language impairment and typical language development: Theoretical and clinical implications. *Developmental Science*, 18, 373–388.
- Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., Jr., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *The Elementary School Journal*, 110, 1–18.

- Dawson, J., Stout, C., Eyer, J. A., Tattersall, P., Fonkalsrud, J., & Croley, K. (2005). *Structured Photographic Expressive Language Test—Preschool 2*. DeKalb, IL: Janelle Publications.
- Deák, G. O., Ray, S. D., & Pick, A. D. (2002). Matching and naming objects by shape or function: Age and context effects in preschool children. *Developmental Psychology*, 38, 503–518.
- Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test—III*. Circle Pines, MN: American Guidance Service.
- Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, 37, 66–108.
- Fey, M. E., Leonard, L. B., Bredin-Oja, S. L., & Deevy, P. (2017). A clinical evaluation of the competing sources of input hypothesis. *Journal of Speech, Language, and Hearing Research*, 60, 104–120.
- Florit, E., Roch, M., & Levorato, M. C. (2011). Listening text comprehension of explicit and implicit information in preschoolers: The role of verbal and inferential skills. *Discourse Processes*, 48, 119–138.
- Gertner, B. L., Rice, M. L., & Hadley, P. A. (1994). Influence of communicative competence on peer preferences in a preschool classroom. *Journal of Speech and Hearing Research*, 37(4), 913–923.
- Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation—Second Edition*. Circle Pines, MN: American Guidance Services.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431–436.
- Gómez, R. L., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, 7, 183–206.
- Graham, S. A., Williams, L. D., & Huber, J. F. (1999). Preschoolers' and adults' reliance on object shape and object function for lexical extension. *Journal of Experimental Child Psychology*, 74(2), 128–151.
- Gray, S. (2004). Word learning by preschoolers with specific language impairment predictors and poor learners. *Journal of Speech, Language, and Hearing Research*, 47, 1117–1132.
- Gray, S. (2005). Word learning by preschoolers with specific language impairment effect of phonological or semantic cues. *Journal of Speech, Language, and Hearing Research*, 48(6), 1452–1467.
- Gray, S., & Brinkley, S. (2011). Fast mapping and word learning by preschoolers with specific language impairment in a supported learning context: Effect of encoding cues, phonotactic probability, and object familiarity. *Journal of Speech, Language, and Hearing Research*, 54, 870–884.
- Gray, S., Pittman, A., & Weinhold, J. (2014). Effect of phonotactic probability and neighborhood density on word-learning configuration by preschoolers with typical development and specific language impairment. *Journal of Speech, Language, and Hearing Research*, 57, 1011–1025.
- Gray, S., Plante, E., Vance, R., & Henriksen, M. (1999). The diagnostic accuracy of four vocabulary tests administered to preschool-age children. *Language, Speech, and Hearing Services in Schools*, 30(2), 196–206.
- Greenslade, K. J., Plante, E., & Vance, R. (2009). The diagnostic accuracy and construct validity of the Structured Photographic Expressive Language Test—Preschool. *Language, Speech, and Hearing Services in Schools*, 40, 150–160.
- Grunow, H., Spaulding, T. J., Gómez, R. L., & Plante, E. (2006). The effects of variation on learning word order rules by adults with and without language-based learning disabilities. *Journal of Communication Disorders*, 39, 158–170.
- Haebig, E., Sterling, A., & Hoover, J. (2016). Examining the language phenotype in children with typical development, specific language impairment, and fragile X syndrome. *Journal of Speech, Language, and Hearing Research*, 59, 1046–1058.
- Hick, R. F., Joseph, K. L., Conti-Ramsden, G., Serratrice, L., & Faragher, B. (2002). Vocabulary profiles of children with specific language impairment. *Child Language Teaching and Therapy*, 18, 165–180.
- Hoover, J. R., Storkel, H. L., & Hogan, T. P. (2010). A cross-sectional comparison of the effects of phonotactic probability and neighborhood density on word learning by preschool children. *Journal of Memory and Language*, 63, 100–116.
- Jones, S. S., & Smith, L. B. (1998). How children name objects with shoes. *Cognitive Development*, 13, 323–334.
- Jones, S. S., & Smith, L. B. (2002). How children know the relevant properties for generalizing object names. *Developmental Science*, 5, 219–232.
- Kamhi, A. G. (2014). Improving clinical practices for children with language and learning disorders. *Language, Speech, and Hearing Services in Schools*, 45, 92–103.
- Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 53, 739–756.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children—Second Edition*. Circle Pines, MN: American Guidance Services.
- Kemler Nelson, D. G., Russell, R., Duke, N., & Jones, K. (2000). Two-year-olds will name artifacts by their functions. *Child Development*, 71, 1271–1288.
- Kim, Y. S., Al Otaiba, S., Puranik, C., Folsom, J. S., & Grulich, L. (2014). The contributions of vocabulary and letter writing automaticity to word reading and spelling for kindergartners. *Reading and Writing*, 27, 237–253.
- McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 45, 998–1014.
- Meyers-Denman, C. N., & Plante, E. (2016). Dose schedule and enhanced conversational recast treatment for children with specific language impairment. *Language, Speech, and Hearing Services in Schools*, 47, 334–346.
- Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, 16, 91–121.
- Oetting, J. B., McDonald, J. L., Seidel, C. M., & Hegarty, M. (2016). Sentence recall by children with SLI across two non-mainstream dialects of English. *Journal of Speech, Language, and Hearing Research*, 59, 183–194.
- Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science*, 21, 1894–1902.
- Plante, E. (1998). Criteria for SLI: The Stark and Tallal legacy and beyond. *Journal of Speech, Language, and Hearing Research*, 41, 951–957.
- Plante, E., Bahl, M., Vance, R., & Gerken, L. A. (2011). Beyond phonotactic frequency: Presentation frequency effects on word productions in specific language impairment. *Journal of Communication Disorders*, 44, 91–102.
- Plante, E., Ogilvie, T., Vance, R., Aguilar, J. M., Dailey, N. S., Meyers, C., . . . Burton, R. (2014). Variability in the language input to children enhances learning in a treatment context. *American Journal of Speech-Language Pathology*, 23, 530–545.

- Rice, M. L., Buhr, J., & Oetting, J. B. (1992). Specific-language-impaired children's quick incidental learning of words the effect of a pause. *Journal of Speech and Hearing Research*, 35, 1040–1048.
- Rice, M. L., Buhr, J. C., & Nemeth, M. (1990). Fast mapping word-learning abilities of language-delayed preschoolers. *Journal of Speech and Hearing Disorders*, 55, 33–42.
- Rice, M. L., & Hoffman, L. (2015). Predicting vocabulary growth in children with and without specific language impairment: A longitudinal study from 2;6 to 21 years of age. *Journal of Speech, Language, and Hearing Research*, 58, 345–359.
- Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research*, 37, 106–122.
- Richtsmeier, P. T., Gerken, L. A., Goffman, L., & Hogan, T. (2009). Statistical frequency in perception affects children's lexical production. *Cognition*, 111, 372–377.
- Snowling, M. J., Duff, F. J., Nash, H. M., & Hulme, C. (2016). Language profiles and literacy outcomes of children with resolving, emerging, or persisting language impairments. *Journal of Child Psychology and Psychiatry*, 57, 1360–1369.
- Suanda, S. H., Mugwanya, N., & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology*, 126, 395–411.
- Suggate, S., Reese, E., Lenhard, W., & Schneider, W. (2014). The relative contributions of vocabulary, decoding, and phonemic awareness to word reading in English versus German. *Reading and Writing*, 27, 1395–1412.
- Torkildsen, J., Dailey, N. S., Aguilar, J. M., Gómez, R., & Plante, E. (2013). Exemplar variability facilitates rapid learning of an otherwise unlearnable grammar by individuals with language-based learning disability. *Journal of Speech, Language, and Hearing Research*, 56, 618–629.
- Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities*, 45, 453–466.
- Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's more like it: Multiple exemplars facilitate word learning. *Infant and Child Development*, 23, 105–122.
- Vlach, H., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology*, 3, 46.
- Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology*, 108, 394–401.
- Vuolo, J., Goffman, L., & Zelaznik, H. N. (2017). Deficits in coordinative bimanual timing precision in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 60, 393–405.
- Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, 13, 70–77.
- Williams, K. T. (2007). *EVT-2: Expressive Vocabulary Test*. San Antonio, TX: Pearson Assessments.
- Wray, C., Norbury, C. F., & Alcock, K. (2016). Gestural abilities of children with specific language impairment. *International Journal of Language & Communication Disorders*, 51, 174–182.