JSLHR

Research Article

The Structure of Word Learning in Young School-Age Children

Shelley Gray,[a] Hope Lancaster,[a] Mary Alt,[b] Tiffany P. Hogan,[c] Samuel Green,[a,†] Roy Levy,[a] and Nelson Cowan[d]

Purpose: We investigated four theoretically based latent variable models of word learning in young school-age children.
Method: One hundred sixty-seven English-speaking second graders with typical development from three U.S. states participated. They completed five different tasks designed to assess children's creation, storage, retrieval, and production of the phonological and semantic representations of novel words and their ability to link those representations. The tasks encompassed the triggering and configuration stages of word learning.
Results: Results showed that a latent variable model with separate phonological and semantic factors and linking indicators constrained to load on the phonological factor best fit the data.
Discussion: The structure of word learning during triggering and configuration reflects separate but related phonological and semantic factors. We did not find evidence for a unidimensional latent variable model of word learning or for separate receptive and expressive word learning factors. In future studies, it will be interesting to determine whether the structure of word learning differs during the engagement stage of word learning when phonological and semantic representations, as well as the links between them, are sufficiently strong to affect other words in the lexicon.

Word learning studies are important for both theoretical and clinical purposes. Theoretically, word learning provides testable hypotheses about a critical component of language acquisition. Clinically, word learning provides the opportunity to observe dynamic learning under well-controlled conditions by participants who bring different abilities to the task. Although we know a good deal about variables that affect word learning, we do not yet have a unifying theoretical model of word learning that defines the factors involved and the relations among factors over time. This is needed because cognitive science has many examples of theoretical models that have helped move science forward, including Baddeley and colleagues' (Baddeley, 2000; Baddeley & Hitch, 1974) and Cowan's (2001) working memory models, Carroll's (1993) model of intelligence, and Gough and colleagues' (Gough & Tunmer, 1986; Hoover & Gough, 1990) Simple View of Reading.

[a]Arizona State University, Tempe
[b]University of Arizona, Tucson
[c]MGH Institute of Health Professions, Boston, MA
[d]University of Missouri, Columbia

Correspondence to Shelley Gray: Shelley.Gray@asu.edu

[†]In memory of our colleague Samuel (Sam) Green, who passed away during the preparation of this article. We gratefully acknowledge his valuable contributions to this research.
Editor-in-Chief: Sean M. Redmond
Editor: Filip Smolik

Received June 1, 2019
Revision received September 7, 2019
Accepted January 22, 2020
https://doi.org/10.1044/2020_JSLHR-19-00186

Theoretical Models of Word Learning

We have theoretical models that explain how very young children solve the problem of initial word learning. Hollich et al. (2000) proposed the "'emergentist coalition model" to explain how children learn their first words in natural contexts. The model considers lexical acquisition to be the simultaneous product of cognitive constraints, social-pragmatic factors, and attentional mechanisms. It addresses principles that first allow children to deduce that words refer to objects or actions in their environment. Like related seminal theories of initial word learning (e.g., Markman, 1989; Merriman & Bowman, 1989; Waxman & Kosowski, 1990) and computational word learning (e.g., Fazly et al., 2010; Frank et al., 2009; Yu & Ballard, 2007), the emergentist coalition model does not consider how word learning proceeds once word learning principles are well established.

Constructivist learning theories such as the social-pragmatic theory proposed by M. Tomasello (2000) do consider word learning beyond initial stages. They emphasize that children construct meaning from their experiences and

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

from reflecting on these experiences. According to this view, word learning is an inherently social process whereby children learn words as they interpret adult intentions within their own cultural context. Accordingly, constructivists emphasize the importance of children being active, creative participants in word learning (Lin, 2015). Similarly, Bloom (2000) emphasizes the complex interaction of conceptual, social, and linguistic processes needed for children to develop a rich mental lexicon. Yet, as important as this work is to our understanding of word learning, the structure of word learning within these theories has not been tested experimentally.

Some authors of word learning research studies refer to word learning models or theories in their work (e.g., Magro et al., 2018); however, these studies typically address a single dimension of word learning such as phonological word forms. For example, Norris et al. (2017) studied 26 young adults (ages 16–24 years) to determine whether a common mechanism underlies phonological word form learning and Hebb learning. Hebb's (1949) rule predicts that connections between two neurons increase in strength if they fire simultaneously. Results suggested that participants' learning of novel word forms (phoneme sequences) followed the same pattern whether the phonemes were presented individually in sequence or in a single nonword. This parallel in learning between discrete sequences of phonemes typical of a Hebb paradigm as well as learning the novel nonword forms led the authors to conclude that a Hebb learning model presents a "viable model for phonological word form learning" (p. 857).

Neurological Models of Word Learning

The lack of theoretical models underlying behavioral measures of word learning extends to neuroscientific models of word learning; however, the "complementary learning systems approach" (McClelland et al., 1995; Shtyrov, 2012) does propose a two-stage process describing how initial phonological aspects of word learning take place quickly in the brain in the hippocampus (Suzuki, 2006). Then, over time, a more gradual process ensues involving interactions between the hippocampus, neocortex, and subcortical structures to form traces of newly learned words in long-term memory (Born et al., 2006; McClelland et al., 1995). Work in this area highlights the importance of sleep in the initial consolidation of new word learning (Davis & Gaskell, 2009; Gaskell & Dumay, 2003), especially in children (Weighall et al., 2017).

Initial efforts by researchers to track word learning in the brain utilized hemodynamic measures to assess blood oxygen level–dependent activation when participants were exposed to novel words (e.g., Breitenstein et al., 2005; Davis & Gaskell, 2009). Researchers were able to document changes in activation in the hippocampus, right inferior frontal gyrus, left fusiform gyrus, and left inferior parietal lobe during the word learning process. Although exciting, a drawback of neuroimaging studies is that they cannot document the temporal resolution of learning and they cannot measure neural processes directly (Shtyrov, 2012). Thus, researchers have also used electroencephalography (EEG; e.g., Shtyrov, 2011), magnetoencephapholography, and event-related potentials to track neural activity during word learning. For example, in a fast mapping EEG study of adults that compared neural activity for known and novel words, Shtyrov et al. (2010) showed that activation for novel versus known words increased after 14 min of exposure to the novel words, which the authors interpreted as evidence for "rapid mapping of new word forms onto neural representations" (p. 16864). Additional experiments showing rapid ability to map word forms with meaning include those by Mestres-Misse et al. (2007), who used event-related potentials to show that, after only three exposures, adults' brain potentials for lexical and semantic processing of novel words in meaningful reading contexts were not distinguishable from real words, and by Batterink and Neville (2011), who showed that, in as little as 10 exposures, novel word representations elicited a robust N400 during lexical decision and word recognition tasks.

Recently, Partanen et al. (2017) utilized magnetoencephapholography to study automatic word form acquisition in 5- to 12-year-old children comparing real and novel words composed of native or nonnative phonology along with nonspeech sounds. They measured brain dynamics as children listened passively to the word stimuli for 20 min while they watched a silent movie. They found distinct spatiotemporal patterns of activation in the brain for native versus nonnative phonology and nonspeech sounds, with the former observed in the left temporal region and the latter in both the right and left hemispheres. The authors interpreted these dynamic changes as evidence for a "rapid …dynamic build-up of memory traces for novel acoustic information in the children's brain" (p. 450). Recently, Abel et al. (2017) used EEG recordings to assess word learning in 11- to 14-year-old children. They found attenuated N400s for words children learned the meaning of versus those they did not and that, once learned, the N400s were similar to those for known words.

These neurological studies of word learning capitalize on real-time measures of brain activation with very interesting results; however, they are not free of potential confounds also inherent in behavioral studies of word learning. Gray et al. (2014) encouraged word learning researchers to consider and carefully describe key factors that affect word learning in their studies, including word, referent, and learner characteristics, the learning context, and the stage of word learning addressed. A unified theoretical model of word learning that spans behavioral and neurological research could promote this goal.

The Word Learning Process

Word learning accrues incrementally across several theoretical stages. Hoover et al. (2010) described the "triggering" stage when a child hears a word and recognizes that it is new, thus triggering attention to and storage of the word. Carey and Bartlett (1978) coined the term "fast mapping” to describe young children’s remarkable ability to hear a word and create initial phonological and semantic representations, as well as links between them, with very few exposures. This is the stage most often studied in neurological studies of word learning. Carey (2010) also discussed the “extended mapping” process that leads to adultlike understanding of the meaning of a word over time. Leach and Samuel (2007) described overlapping “configuration” and “engagement” stages of word learning, which correspond to Carey’s extended mapping. They proposed that configuration encompasses learning of the phonological and/or orthographic forms of a word, the meaning of the word and its syntactic roles, plus linking these forms. Leach and Samuel proposed that engagement occurs when a new word affects other words in the lexicon. Neighborhood density effects demonstrate this when phonologically similar words compete for retrieval. Recently, Weighall et al. (2017) used an eye-tracking experiment to show that phonological competition may occur in children relatively quickly (same day) but is increased following a period of sleep. They concluded that competition effects are stronger for existing versus newly learned words in children, but also that, “different aspects of new word learning follow different time courses” (p. 13).

Each of these stages encompasses the creation and/or enrichment of newly formed phonological and semantic representations; thus, phonology and semantics form the foundation of word learning success. When reading and writing are involved, orthographic (letter pattern) representations also come into play. To produce words accurately, detailed phoneme-by-phoneme representations are required, along with an articulatory representation of the word called the “articulatory score” (Indefrey, 2011). These orthographic and articulatory representations depend on well-specified phonological representations.

Structural Models of Word Learning

To date, we have no empirical tests of a comprehensive word learning model. This will require considerable resources because of the large number of participants needed to evaluate structural equation models and because the engagement stage of word learning, where a newly learned word influences other words in the lexicon, will require multiple probes over time. Nevertheless, our long-term goal is to test empirically based models of the entire word learning process; but as a first step, we tested a latent variable model of the triggering and configuration stages of word learning in young school-age children. Because phonological and semantic representations are central to each stage of word learning, we included multiple tasks assessing newly formed phonological and semantic representations of words. We used structural equation modeling for our analyses because this allowed us to test hypotheses about the complex, multidimensional relationships among our observed and latent variables (Hoyle, 1995). Observed variables are those directly measured by the researcher such as the number of new words produced. In contrast, latent variables cannot be directly observed but rather are hypothetical constructs inferred from responses to observed variables (MacCallum & Austin, 2000). For example, if a set of measures assessed a child’s understanding of a word from different perspectives, such as their ability to point to the referent or to give a synonym, and the scores on these measures turned out to be highly correlated, this would provide evidence for an underlying construct of “receptive language.” The receptive language factor would be assumed to cause the correlations among scores on the word understanding tasks.

In addition to their usefulness in making underlying constructs explicit, latent variables also have the advantage of having no associated measurement error because they are not direct measures of a behavior. This allows examination of common variance and provides the opportunity for researchers to answer interrelated questions using a single, comprehensive analysis (but see Tarka, 2018, for a discussion of opportunities and threats associated with structural equation modeling).

To determine which latent variable models of the early stages of word learning to test, we relied on published word learning studies and vocabulary research. Although there is no model of word learning for children, existing studies provide insight into factors that should be considered in a word learning model. We tested four potential models: unidimensional, receptive/expressive, phonological/semantic, and a three-factor model representing the creation, linkage, and retrieval of new words.

Our first latent variable model was unidimensional in nature. We hypothesized that, in second-grade children, the ability to learn new words might depend on a single underlying language learning factor. Recent studies of oral language including vocabulary, grammar, and listening comprehension suggest that, in young children, oral language may be a single construct (Language and Reading Research Consortium [LARRC], 2017). LARRC administered receptive and expressive vocabulary and grammar measures and listening comprehension measures to 1,869 children in preschool through third grade. At each grade level, multiple measures of oral language and listening comprehension loaded on separate factors, but the factors were highly correlated at .91, suggesting they were not independent. Similar findings were reported by Anthony et al. (2014), Bornstein et al. (2014), Foorman et al. (2015), LARRC (2015), Lonigan and Milburn (2017), and Tomblin and Zhang (2006). These studies identified one or two factors for oral language, but when two factors were identified (vocabulary and grammar), they were highly correlated.

Although investigations of the dimensionality of language have found no evidence that receptive and expressive vocabulary are separate factors, clinicians and researchers often assess each separately. This may be due to the availability of separate receptive and expressive vocabulary tests, or to observations that it is easier to recognize or comprehend a word than to produce it. A 2010 meta-analysis of vocabulary interventions did not find differences in treatment effects for receptive versus expressive measures (Marulis & Neuman's, 2010), but a 2010 meta-analysis of word learning studies comparing children with typical development (TD) to those with primary language impairment (LI) found larger between-groups differences on receptive/recognition measures than expressive measures (Kan & Windsor, 2010). This research leads to the question of whether receptive and expressive word learning measures tap different constructs or the same underlying construct. We tested this in our second latent variable model with receptive word learning as one factor and expressive word learning as the second. Because recent research suggests that receptive and expressive vocabularies are not separate factors, we hypothesized that this model would not provide a good fit to the data.

Our third latent variable model contained phonological and semantic factors. As noted earlier, word learning requires both phonological (word form) and semantic (word referent) skills. Research shows that auditory word form processing primarily involves the superior temporal gyrus (Booth et al., 2002a, 2002b; Booth, Burman, Meyer, Gitelman, et al., 2003; Booth, Burman, Meyer, Zhang, et al., 2003) but that the brain stores semantic information in distributed patterns of related concepts throughout the brain (Huth et al., 2016; R. Tomasello et al., 2017). Based on evidence that phonological and semantic word learning processes occur in different areas of the brain, we tested out third latent variable model with separate phonological and semantic factors. We hypothesized that, because of robust neurological and behavioral evidence showing the importance of phonological and semantic representations of words, our data would fit this model well.

Our fourth latent variable model included factors representing the word learning process: creation and storage of new phonological and semantic representations, linking those representations, as well as the retrieval, recreation, and production of words when children were asked to produce words. We hypothesized that factors related to the creation and storage of new phonological and semantic representations could be distinct from those required to link representations or to produce words, and thus, the data could fit this model well. For example, Booth et al. (2004) reported that, when phonological and semantic representations interact, mediation occurs in the supramarginal and angular gyri, engaging different brain areas from processing associated with creating and storing phonological and semantic representations. Thus, our fourth and final latent variable model tested the word learning process as a whole.

The strengths of this study are that it employed a wide variety of word learning tasks to test four plausible latent variable models of word learning during the triggering and configuration stages of word learning in children who, by virtue of their age and experience, have already integrated initial principles of word learning into their language learning repertoire. In addition, we carefully controlled variables known to affect word learning including phonotactic probability, neighborhood density, referent characteristics, number of exposures, word learning context, and learner characteristics (Gray et al., 2014).

Of the four models tested (unidimensional, receptive/expressive, phonological/semantic, and create/recreate/link), we hypothesized that the latter two were most likely to fit the data well because it is well known that both phonological and semantic representations are necessary to establish words in the lexicon and because the link between those representations must be established for either representation to activate the other in long-term memory.

Method

This research was approved by the internal review boards of Arizona State University and The University of Arizona where data were collected. Procedures adhered to ethical standards for research conducted with human subjects. Parents gave their consent for children to participate in the study, and children gave their assent.

Second-grade children with TD from rural and metropolitan areas of Arizona participated in this study. We enrolled 167 children who were part of a larger study on working memory and word learning. There were 72 girls and 95 boys. For ethnicity, 87% reported non-Hispanic, 12% reported Hispanic, and 1% provided no report. For race, 2% reported American Indian or Alaska Native, 2% reported Asian, 2% reported Black, 81% reported White, 12% reported more than one race, and 1% did not report. Table 1 provides additional descriptive information about the participants.

Inclusionary criteria included (a) passing a bilateral hearing screening, (b) passing a color vision screening, (c) passing a near-vision acuity screening, (d) enrolled in or just completed second grade, (e) no history of neuropsychiatric disorders (e.g., attention-deficit/hyperactivity disorder [ADHD], autism spectrum disorder) by parent report, (f) spoke monolingual English by parent report,

Table 1. Participant characteristics and test scores.

Measure	*M*	*SD*
Age in months	92.82	4.97
Mother's education in years	15.39	1.66
GFTA-2 Articulation Accuracy percentile	50.89	8.54
K-ABC2 Nonverbal Index standard score	117.60	15.53
TOWRE-2 Word/Nonword standard score	109.45	8.40
CELF-4 Core Language standard score	108.75	9.58
EVT-2 standard score	112.39	10.95
WRMT III-PC standard score	108.23	9.85
ADHD Rating Scale-IV raw score	10.19	8.77

Note. GFTA-2 = Goldman-Fristoe Test of Articulation–Second Edition (Goldman & Fristoe, 2000); K-ABC2 = Kaufman Assessment Battery for Children, Second Edition (Kaufman & Kaufman, 2004); TOWRE-2 = Test of Word Reading Efficiency–Second Edition (Torgesen et al., 2012); CELF-4 = Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel et al., 2003); EVT-2 = Expressive Vocabulary Test–Second Edition (Williams, 2007); WRMT III-PC = Woodcock Reading Mastery Test–Third Edition Passage Comprehension (Woodcock, 2011); ADHD Rating Scale-IV = ADHD Rating Scale–Fourth Edition: Home Version (DuPaul et al., 1998).

(g) standard score of ≥ 75 on the Nonverbal Index of the Kaufman Assessment Battery for Children, Second Edition (Kaufman & Kaufman, 2004), (h) no history of special education services or repeating a grade, (i) standard score of > 30th percentile on the Goldman-Fristoe Test of Articulation–Second Edition (Goldman & Fristoe, 2000) unless scores below that percentile were due to consonant errors on a single sound, (j) standard score of > 87 on the core language composite of the Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel et al., 2003), and (i) second-grade composite standard score of > 95 on the Test of Word Reading Efficiency–Second Edition (Torgesen et al., 2012).

We administered an 18-item ADHD Rating Scale–Fourth Edition (Home Version; DuPaul et al., 1998) that asked parents to rate their child's behavior over the past 6 months. The scale items were adapted from the *Diagnostic and Statistical Manual of Mental Disorders–IV–Text Revision* (American Psychiatric Association, 2000) diagnostic criteria for ADHD. The highest possible score was 54, which would indicate a high level of concern about attention and/or hyperactivity. Children with a diagnosis of ADHD were excluded from the study, but we measured functional attention for descriptive purposes. We also report standard scores on the Woodcock Reading Mastery Test–Third Edition for the Passage Comprehension subtest (Woodcock, 2011). Participant characteristics and test scores are reported in Table 1.

General Procedures

Trained research assistants (primarily retired teachers or college students) administered assessments and experimental measures individually in a quiet room at the child's school, a local library, our laboratory, or the child's home. Research assistants were required to pass a quiz and two fidelity checks demonstrating their ability to administer and score each assessment correctly and, for the computer-based word learning experiments, to set up the equipment correctly, to guide the child through the automated word learning tasks, to provide correct feedback if the child asked questions, and to complete forms such as hearing screening and child assent.

The experimental tasks are from the Comprehensive Assessment Battery for Children–Word Learning (Gray et al., 2020). We designed this battery to test phonological and semantic aspects of word learning in multiple ways during the triggering and configuration stages of word learning. We included nouns and verbs because, in young children, nouns appear easier to learn than verbs (Bornstein et al., 2004; Childers & Tomasello, 2006; Gentner, 1982, 2006; Maguire et al., 2005). As described in Table 2, we directly manipulated phonological and semantic aspects of word learning (e.g., word and referent characteristics) because one purpose of the battery is to identify the source of word learning difficulties in children. That was not the focus of this study, but these manipulations also meet another purpose of the study, to yield multiple indicators of word learning necessary to test structural equation models.

In addition to the word learning tasks reported in this article, children completed more word learning and working memory tasks.[1] All were presented in a computer-based, pirate-themed game that took approximately six 2-hr sessions to complete over a period of about 2 weeks. The five word learning games reported in this study each taught four different nonwords and took about 30 min per game to complete. Children played only one word learning game per day. One game manipulated word length using nouns, one phonological similarity using nouns, one location of the referent (stationary vs. changing position) using nouns, one visual similarity of the referents using nouns, and one different actions using verbs. Each game was presented on a different day with the order randomized by the computer. Children earned virtual coins as they played to spend on their pirate at a virtual pirate store. Children were seated in front of a touch screen computer monitor beside a trained research assistant. The child and research assistant wore headsets with integrated microphones used to record children's verbal responses for later transcription in the lab.

Materials and Tasks

Table 2 provides a description of the word learning stimuli, tasks, and manipulations of the stimuli including an overview of the nonwords and referents, the word learning processes assessed, the experimental manipulations of the stimuli, the type of working memory assessed by each task, and descriptions of the assessment tasks.

Nonwords

We created a pool of low phonotactic probability two-syllable (e.g., /ka mjeg/) and four-syllable (e.g., w^gtifhektUd) consonant–vowel–consonant syllable structure nonwords so that their (a) duration in milliseconds, (b) biphone frequency, and (c) summed biphone probability were very similar. Four 2-syllable nonwords from the pool were randomly assigned to each game (except the game that manipulated word length where 2 two-syllable and 2 four-syllable words were randomly assigned). The nonwords had no phonological neighbors. Nonwords used as verbs were intransitive. A detailed description of the word characteristics may be found in Alt, Arizmendi, et al. (2019). During each word learning game, the computer randomly assigned nonwords to referents for each child.

[1]Participants in this study represent a portion of the participants in a larger sample from the Profiles of Working Memory and Word Learning (POWWER) project funded by National Institutes of Health - National Institute on Deafness and Other Communication Disorders (NIDCD) Grant R01 DC010784. The POWWER project includes the group reported, as well as children with LI, children with dyslexia, and children with comorbid dyslexia and LI. POWWER participants completed a total of six word learning games and a comprehensive battery of working memory tasks (see Cabbage et al., 2017) over the course of at least 6 days. Results of other word learning studies may be found in Alt, Arizmendi, et al. (2019); Alt, Gray, et al. (2019); Alt et al. (2017); Baron et al. (2018); Erikson et al. (2018).

Table 2. Description of word learning stimuli, tasks, and manipulations.

Stimuli	Process assessed	Experimental manipulation	Type of working memory assessed	Assessment task
Noun nonwords CVC–CVC wo-syllable structure; no phonological neighbors (low neighborhood density); low biphone phonotactic probability (1.0039–1.009)	Create and store phonological form (receptive)	2 vs. 4 syllables Phonologically similar vs. phonologically dissimilar words	Phonological loop capacity (length) Specificity of stored phonological representation	Mispronunciation detection A monster appears on the screen, and the child hears either the correct name or a foil. The child presses a key for “yes” if correct name or “no” if incorrect name. They receive immediate feedback on whether they responded correctly.
	Retrieve and produce phonological form (expressive)	2 vs. 4 syllables Phonologically similar vs. phonologically dissimilar	Phonological loop capacity (length) Specificity of stored phonological representation	Naming A monster appears on the screen, and the child is asked to name it. Their response is recorded for later scoring. They receive positive feedback for responding, but no feedback on whether their response was correct.
Noun referents Virtual sea monsters all the same size, but varied body shapes, colors, limb shapes, head coverings, and facial features	Create and store semantic representation (receptive)	Stationary referent vs. referent changes location Visually similar referent vs. visually dissimilar referent	Spatial memory Specificity of stored semantic (visual) representation	Visual difference decision A monster appears on the screen. The child is asked to press a key for “yes” if the monster shown is an accurate depiction of one of the learned monsters or press a key for “no” if it is not one of the monsters they have learned. They receive immediate feedback on whether they responded correctly.
	Retrieve and recreate semantic representation (expressive)	Stationary referent vs. referent changes location Visually similar referent vs. visually dissimilar referent	Spatial memory Specificity of stored semantic representation	Visual feature recall The outline of a monster appears on the screen along with a menu that includes choices of monster colors, eyes, arms, and head coverings. The child is asked to choose the correct features for that monster and drag them onto the monster. They receive immediate feedback based on the number of correct selections they made.
	Link phonological form and semantic representation (link)	2 vs. 4 syllables Phonologically similar vs. phonologically dissimilar words Stationary referent vs. referent changes location Visually similar referent vs. visually dissimilar referent	Phonological loop capacity (length) Specificity of stored phonological representation Spatial memory Specificity of stored semantic representation	Phonological–visual linking Four monsters appear on the screen. The child hears the name of one monster and is asked to choose the monster that goes with the name. They receive immediate feedback on whether they responded correctly.
Verb nonwords CVC–CVC two-syllable structure; no phonological neighbors (low neighborhood density); low biphone phonotactic probability (1.0039–1.009)	Create and store phonological form (receptive)	None	Specificity of stored phonological representation	Mispronunciation detection A monster who is performing an action appears on the screen. The child hears a nonword that is a command for performing an action. The child presses a key for “yes” if the command they hear is correct for the action, or a key for “no” if it is not the correct command for that action. They receive immediate feedback on whether they responded correctly.

(*table continues*)

Table 2. *(Continued).*

Stimuli	Process assessed	Experimental manipulation	Type of working memory assessed	Assessment task
	Retrieve and produce phonological form (expressive)	None	Specificity of stored phonological representation	Naming A monster who is performing an action appears on the screen. The child is asked to say the command for that action. Their response is recorded for later scoring. They receive positive feedback for responding, but no feedback on whether their response was correct.
Verb referents Single virtual sea monster with movement varied by speed, direction, nature of movement, and special effects such as glowing or pulsating	Create and store semantic representation (receptive)	Four different referent actions	Spatial memory Specificity of stored semantic representation	Visual difference decision A monster who is performing an action appears on the screen. The child is asked to press a key for "yes" if the action shown is an accurate depiction of one of the learned actions or press a key for "no" if it is not a learned action. They receive immediate feedback on whether they responded correctly.
	Retrieve and recreate semantic representation (expressive)	Four different referent actions	Spatial memory Specificity of stored semantic representation	Visual feature recall The outline of a monster appears on the screen along with a menu that includes choices of speed, direction, type of movement, and special effects (such as glowing). The child is asked to choose the correct features for that monster and drag them onto the monster. They receive immediate feedback based on the number of correct selections they made.
	Link phonological form and semantic representation (link)	All of the above	All of the above	Phonological–visual linking Four different monsters appear on the screen. The child hears a name (or action command for verbs). The child is asked to choose which monster (or action for verbs) was correct. They received immediate feedback on whether they responded correctly.

Note. CVC = consonant–vowel–consonant.

Referents

We created a set of colored sea monster drawings to use as referents (see Figure 1 for examples). The monsters differed in shape, color, arm style, eye shape, and type of head covering but were the same overall size. A different set of monsters was used for each task.

Word Learning Procedures

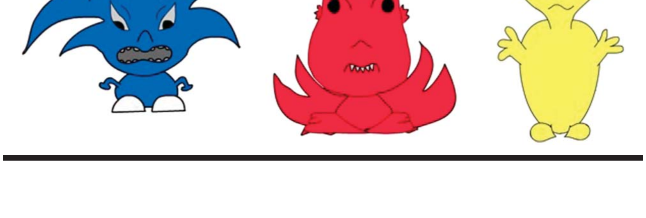
We measured word learning using five different tasks designed to assess children's creation, storage, retrieval, and production of the phonological and semantic (visual) representations of words and their ability to link those representations. Although each word learning game (described in Table 2 and below) featured a different experimental manipulation, the procedures were the same across games. This allowed us to assess the effects of the experimental manipulation on all other aspects of word learning.

Each word learning game described in Table 2 included four blocks. As shown in Table 3, Block 1 assessed fast mapping (including triggering) by providing two exposures to each of the four nonwords. Blocks 2, 3, and 4 assessed configuration by presenting 15 additional exposures per block for the same four words.

At the beginning of each block, the child saw pictures of four sea monsters on the screen and heard a name (or action in the case of verbs). As shown in Figure 2, they were asked to touch the monster that went with the name (or action). They received a gold coin for each correct answer and a rock for each incorrect answer. This phonological–visual linking task assessed children's ability to link new phonological (label) and visual (referent) information for each monster. After completing the linking task, the computer administered the four remaining tasks in random order.

For the mispronunciation detection task, each of four sea monsters came on the screen, one at a time, and the child heard the correct name (or action) for the monster or a unique phonologically related foil with a different final consonant. Children pressed a key to indicate whether the name (action) they heard was correct. They received immediate feedback with a coin or rock. For the naming task, each of the four monsters appeared on the screen, one at a time, and the child was asked to name the monster or the action a monster completed. The computer recorded the child's response for later transcription in the lab. Children received gold coins for attempting to name the monster (action) but no feedback on the accuracy of their response.

Figure 1. Examples of sea monsters used in word learning games.



For the visual difference detection task designed to assess children's semantic representations, a monster appeared on the screen that was the same target monster children had been learning to name or a visually related foil. The foils could vary in one to three ways from the target monster—by color, type of head covering, or eye shape. Children pressed a key to indicate whether the monster was the correct target monster. They received immediate feedback with a gold coin or rock. For the visual feature recall task (see Figure 3), children saw a line drawing outline of a monster beside a visual menu of semantic features including four choices of color, eye shapes, arms, and types of head coverings. They selected one of each feature to put on their monster. When they were satisfied that their feature choice was accurate, they pressed an "I'm done" button. They received a gold coin or rock for each selected semantic feature.

Analytic Approach

We used a series of confirmatory factor analyses (CFAs) to test the structure of novel word learning. By using CFA, we could evaluate several different structures, allowing us to test our hypotheses about the nature of word learning, proposed latent factors, and to estimate interfactor correlations. Consistent with our previous work on understanding the dimensionality of language, a set of four predetermined latent variable models were compared for quality of fit. First, we evaluated a unifactor model of word learning (Model 1; see Figure 4) that tested the hypothesis that, in second-grade children, the ability to learn new words might depend on a single underlying language learning factor. Second, we tested a two-factor model differentiating receptive and expressive forms (Model 2; see Figure 5) based on the clinical notion that receptive and expressive vocabulary represent different underlying constructs. Third, based on the commonly accepted understanding that word learning involves both phonological and semantic processes and evidence that phonological and semantic word learning processes occur in different areas of the brain, we tested a two-factor model differentiating phonological and semantic forms (Model 3; see Figure 7 for a refined version, further described in the next section). Finally, based on research supporting word learning processes during the triggering and configuration stages of word learning, we tested a three-factor model of word learning (create/store, link, and retrieve/recreate/produce: Model 4; see Figure 6). We used the same variables for all latent variable models but with different specifications.

Table 4 presents the means and standard deviations for the word learning variables used in the latent variable models. Table 5 presents the first-order Pearson correlations among the measures, which ranged from small to large. Not all correlations were significant, and it is worth noting that the phonological–visual linking tasks were highly correlated. Table 6 presents the latent factor specification for each variable in the a priori models except the unifactor model. Each model fit accounted for experimenter manipulation as shown by correlated error terms in the models (e.g., correlated

Table 3. Tasks administered in each experimental block.

Task	Block 1	Block 2	Block 3	Block 4
Administered 1st	Phonological–visual linking task 4 words × 2 exposures each	Phonological–visual linking task Same 4 words × 15 exposures each	Phonological–visual linking task Same 4 words × 15 exposures each	Phonological–visual linking task Same 4 words × 15 exposures each
Administered in random order	Mispronunciation detection task Naming task Visual difference detection task Visual feature recall task	Mispronunciation detection task Naming task Visual difference detection task Visual feature recall task	Mispronunciation detection task Naming task Visual difference detection task Visual feature recall task	Mispronunciation detection task Naming task Visual difference detection task Visual feature recall task

error term between Naming Nouns 2 syllables and Naming Nouns 4 syllables). For simplicity, correlated error terms are shown only on Model 1 but apply to each latent variable model.

We conducted the CFAs in RStudio Team (2016) with the lavaan package (Rosseel, 2012) using maximum likelihood parameter estimation with standard errors to handle nonnormality distribution and full information maximum likelihood to handle missing data. We conducted the little test for missing data completely at random and did not reject the null hypothesis of missing data completely at random. Of the 167 children enrolled, from 159 to 162 completed each task as shown in Table 4. Missing data were primarily due to technology failures. Current practice is to use several model fit criteria instead of relying on a single measure. We assessed model fit using a combination of absolute, parsimony, and comparative indices of model fit (Byrne, 2012). Our index of absolute fit was standardized root-mean residual (SRMR), which represents the squared difference between observed and predicted correlations and for which values of < .08 are considered acceptable. Our indices of parsimony included the root-mean-square error of approximation (RMSEA) and the Akaike information criteria (AIC). RMSEA ranges from 0 to 1 with values of < 0.08 representing acceptable fit and values < 0.05 representing close fit (Browne & Cudeck 1993). We report the 90% confidence interval for the RMSEA and the p value for the closeness of fit test, which tests the null hypothesis that RMSEA is ≤ 0.05. For AIC, which is often used when comparing nonnested models, smaller values indicate better fit. We used the Tucker–Lewis Index (TLI) to assess comparative/incremental fit. The TLI is a nonnormed fit index that is analogous to R^2 with a penalty for added parameters. Like R^2, higher values indicate better fit, with the traditional cutoff value for good fit at 0.90. When comparing nested models, we report the Satorra–Bentler rescaled χ^2 (S-B χ^2) difference test. A statistical chi-square difference test indicates that the more complex model fits statistically significantly better than the more restricted or more parsimonious model.

The model fit indices discussed so far can be affected by design features, including sample size (Byrne, 2012). Thus, we want to emphasize the importance of viewing the model fit indices as a set and to pair these statistics with theory and interpretability. Rather than selecting the "best" fitting model based on fit indices alone, we also considered the degree to which constructs are separated and paths fit.

Figure 2. Example of phonological–visual linking task where child makes correct response.

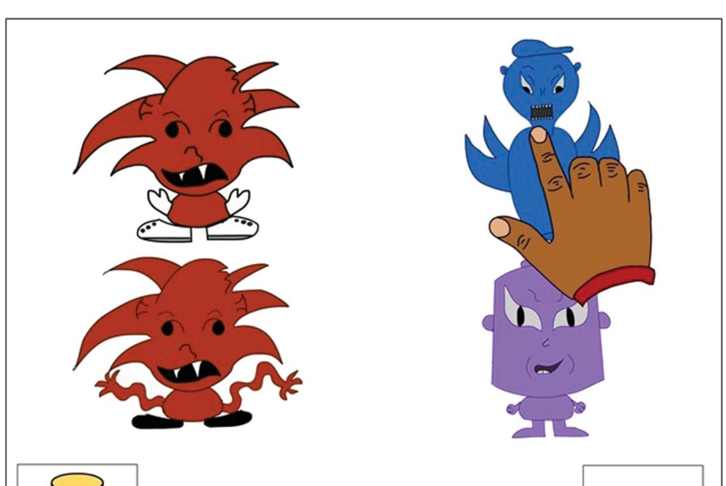
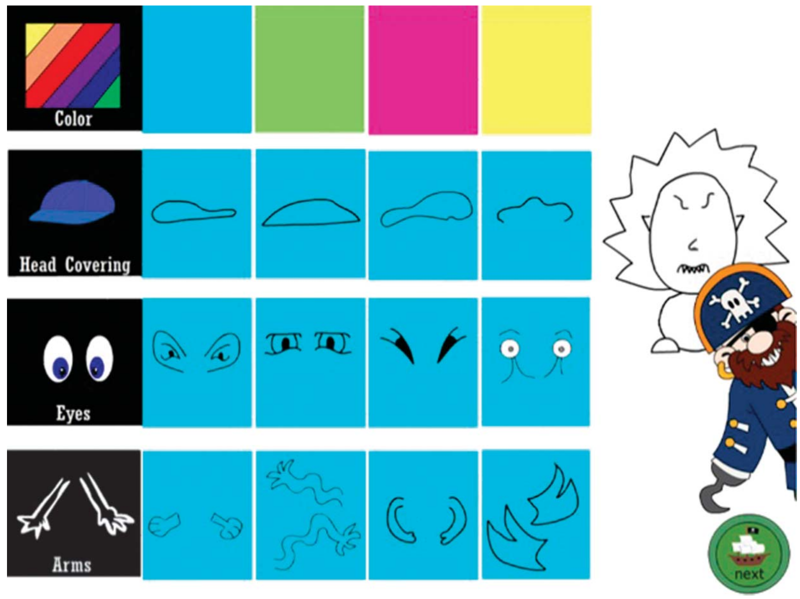


Figure 3. Example of visual feature recall task before child has made any responses.



Therefore, we examined magnitude of the interfactor correlation and considered the average variance extracted (AVE; Hair et al., 2014). The AVE is calculated by squaring and then averaging the standardized loadings of a construct. If the proportion of variance extracted that is unique to a construct (i.e., AVE) is less than the proportion of shared variance between two factors (i.e., squared factor correlation), the evidence for distinct factors is weak (Hair et al., 2014). Additionally, we examined standardized and unstandardized parameter estimates, standardized residual covariances, modification indices, and R^2 values for manifest paths to assess whether the models were meaningful and interpretable. We used path fit to explore model refinement if necessary. Figures 4–7 illustrate the models we present in this study.

Results

Descriptive Statistics and Correlations

Before conducting analyses, we examined the distributions of all measures to check for deviations from normality using histograms, skewness, and kurtosis. No severe departures from normality were observed as none of the skew or kurtosis values were outside the $\pm$ 2 recommendation; however, visual inspection of the histograms showed slight nonnormality. To adjust for slight nonnormality of the data, all analyses conducted used maximum likelihood estimation with robust standard errors.

Table 7 presents model fit statistics for the confirmatory models. For all latent variable models, the following task pairs (illustrated in Figure 4 only) had nonsignificant correlated error terms: Naming Nouns 2 and 4 Syllables, Mispronunciation Detection Nouns Similar and Dissimilar, Visual Difference Decision Nouns Stationary and Moving, and Visual Difference Decision Nouns Visually Similar and Dissimilar. We allowed correlated error terms in the model for tasks that shared the same method (e.g., the only difference between naming two- and four-syllable nouns was that the words differed in syllable length). Because of their common method of measurement, we hypothesized that the variances of the two tasks would overlap.

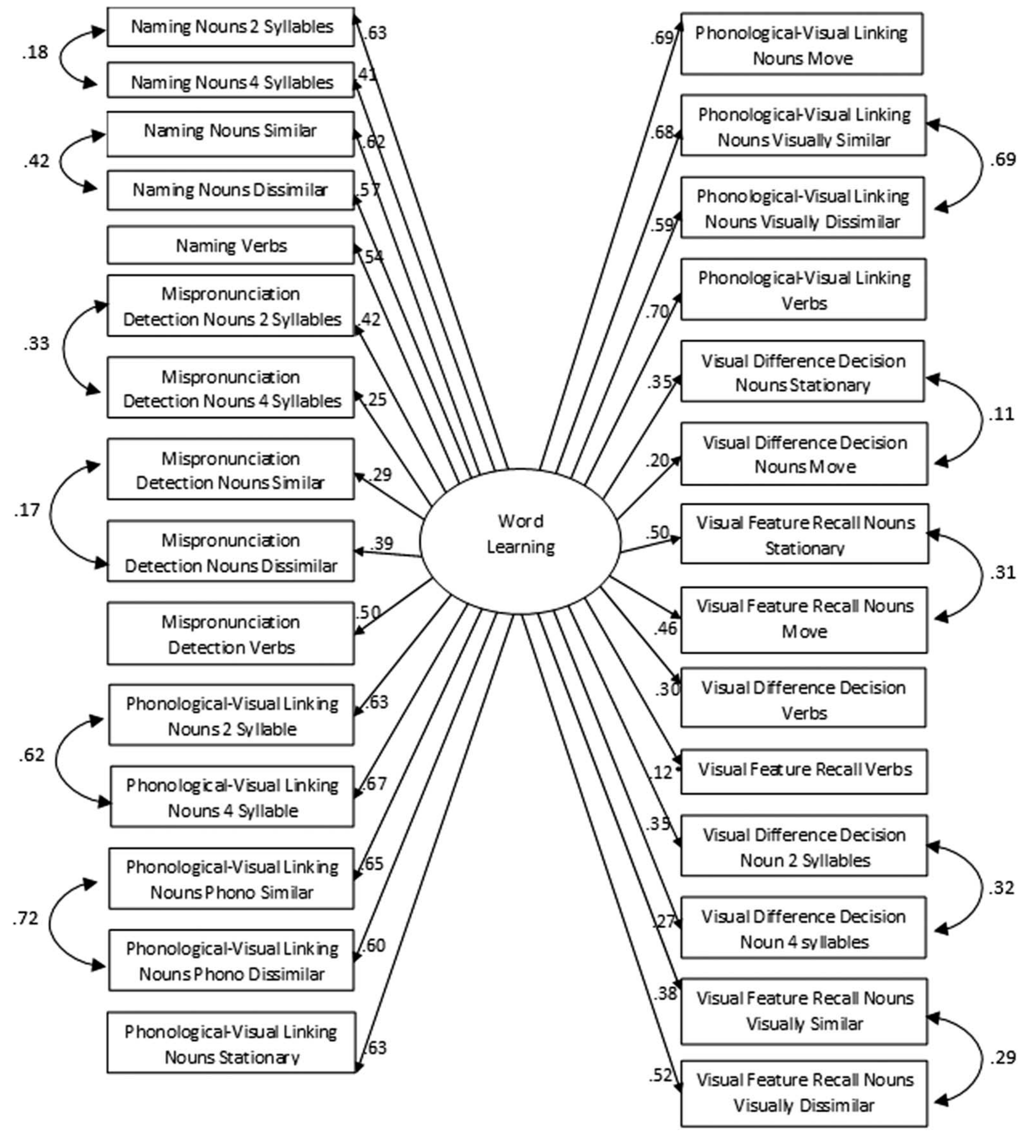
Step 1: Fitting Proposed Models

We first ran a base model (Model 1; see Figure 4) with a single word learning factor. For this unifactor model, the majority of standardized loadings were significant, ranging in size from 0.123 to 0.703, with eight paths less than 0.40 and one nonsignificant path (Visual Feature Recall Verbs). Model fit was within the acceptable range for RMSEA (0.053) and SRMR (0.071), but not for TLI (0.887).

For the receptive–expressive model (Model 2; see Figure 5), fit indices were within the acceptable range. The correlation between the receptive and expressive factors, however, was large and significant (r = .899, p < .001), indicating that these two factors were not separate constructs. Most standardized loadings were significant, and values ranged from 0.117 to 0.705, with nine paths less than 0.40. The S-B χ^2 difference test comparing the nested unifactor model and the two-factor model suggested that there was no difference between these models (see Table 7).

For the phonological–semantic model (Model 3; original version not shown in a figure; refined version shown in Figure 7), fit indices were within the acceptable range. In addition to the correlated error terms shown on Model 1, the phonological–semantic model had nonsignificant correlated errors between the Visual Feature Recall Nouns–Visually Similar and Dissimilar (.29). The correlation between

Figure 4. Base word learning model (Model 1).

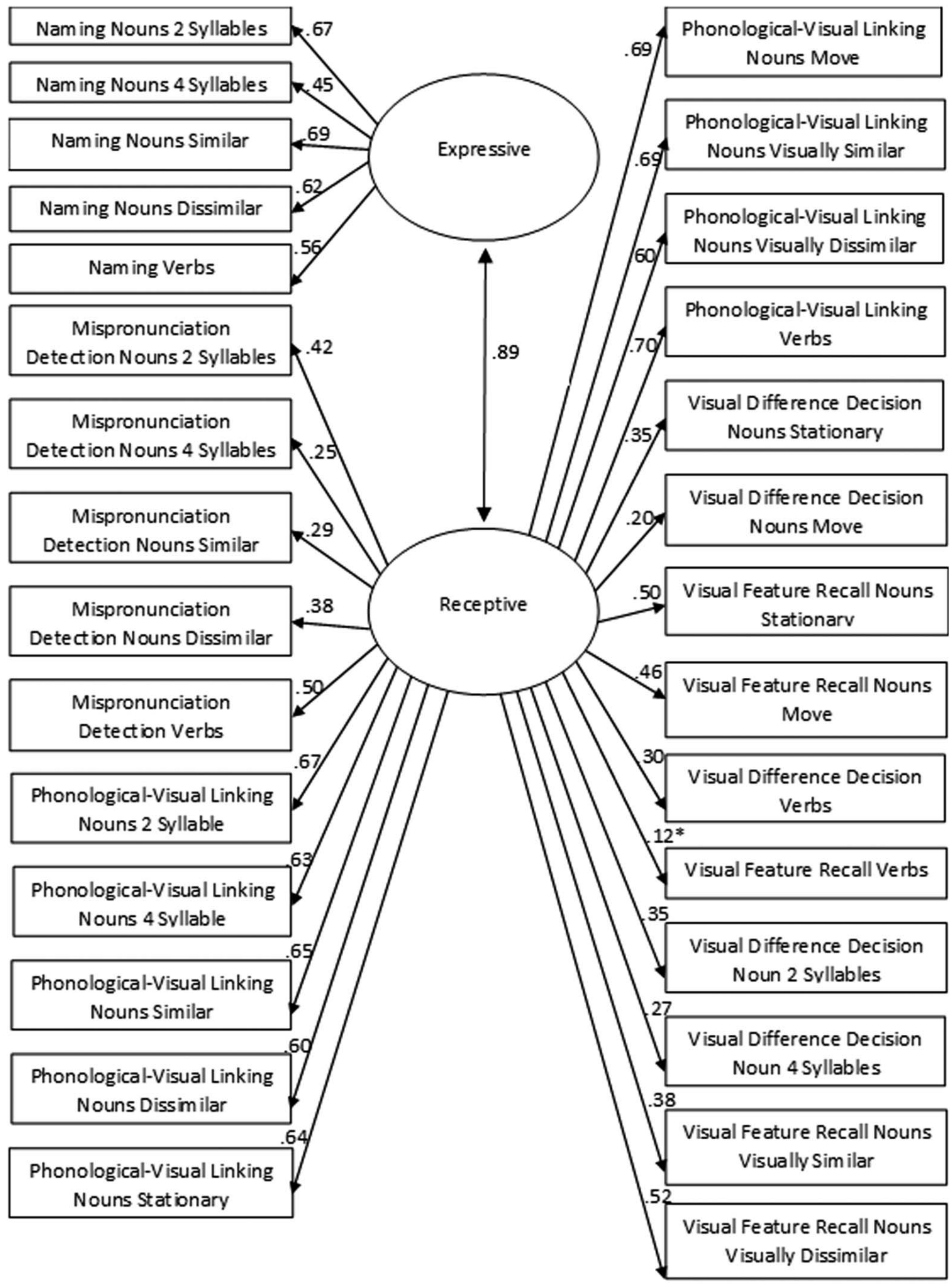


factors was high ($r = .793$, $p < .001$), which indicated that the factors were not completely independent. For the phonological factor, all standardized loadings were significant and ranged from 0.241 to 1.076; in contrast, for the semantic factor, seven out of 16 loadings were not significant and values ranged from −0.026 to 0.664. These loading values indicated that cross-loading the phonological–visual linking tasks was not appropriate. Regardless of the loadings, the S-B χ^2 difference test indicated that the phonological–semantic model was a better fit to the data than the unifactor model.

For the create/store, link, retrieve/recreate/produce model (Model 4; see Figure 6), fit indices were acceptable. All standardized loadings for the create (range: 0.253–0.559) and link (range: 0.633–0.744) factors were significant; however, one loading was not significant for the recreate factor (Visual Feature Recall Verbs, $B = 0.174$, $z = 1.940$, $p = .052$), and values ranged from 0.166 to 0.671. The correlations between latent factors were significant (create/store to recreate/produce, $r = .865$, $p < .001$; create/store to link, $r = .703$, $p < .001$; recreate/produce to link, $r = .851$, $p < .001$). The S-B χ^2 difference test indicated that the create–recreate–link model was a better fit than the unifactor model (Model 1; see Table 7).

In summary, in examining hypothesized structures of novel word learning, we fit a series of four latent variable models, a unifactor model, a receptive–expressive model, a phonological–semantic model, and a three-factor create–recreate–link model. Among these four models, two models had acceptable goodness of fit: phonological–semantic and create–recreate–link. Of these two models, fit was slightly better for the phonological–semantic model.

Figure 5. Receptive and expressive word learning model (Model 2). *Indicates nonsignificant path loading.



However, goodness-of-fit statistics are not the only consideration when selecting models. We must also consider the distinctness of latent factors and model parsimony. Inspection of the manifest paths showed that both the phonological–semantic and create–recreate–link models had several misspecified paths. For the phonological–semantic model, the misspecified paths were those cross-loaded on two factors. In contrast, for the create–recreate–link model, the misspecified paths were across all latent factors. Examination of the covariance between the latent factors in the two models showed that the latent factors were more distinct in the phonological–semantic latent variable model. Thus, based on goodness of fit, manifest path fit, and distinction of latent factors, it was clear that the phonological–semantic model was the best candidate for model refinement.

Figure 6. Create/store, link, retrieve/recreate/produce word learning model (Model 4).

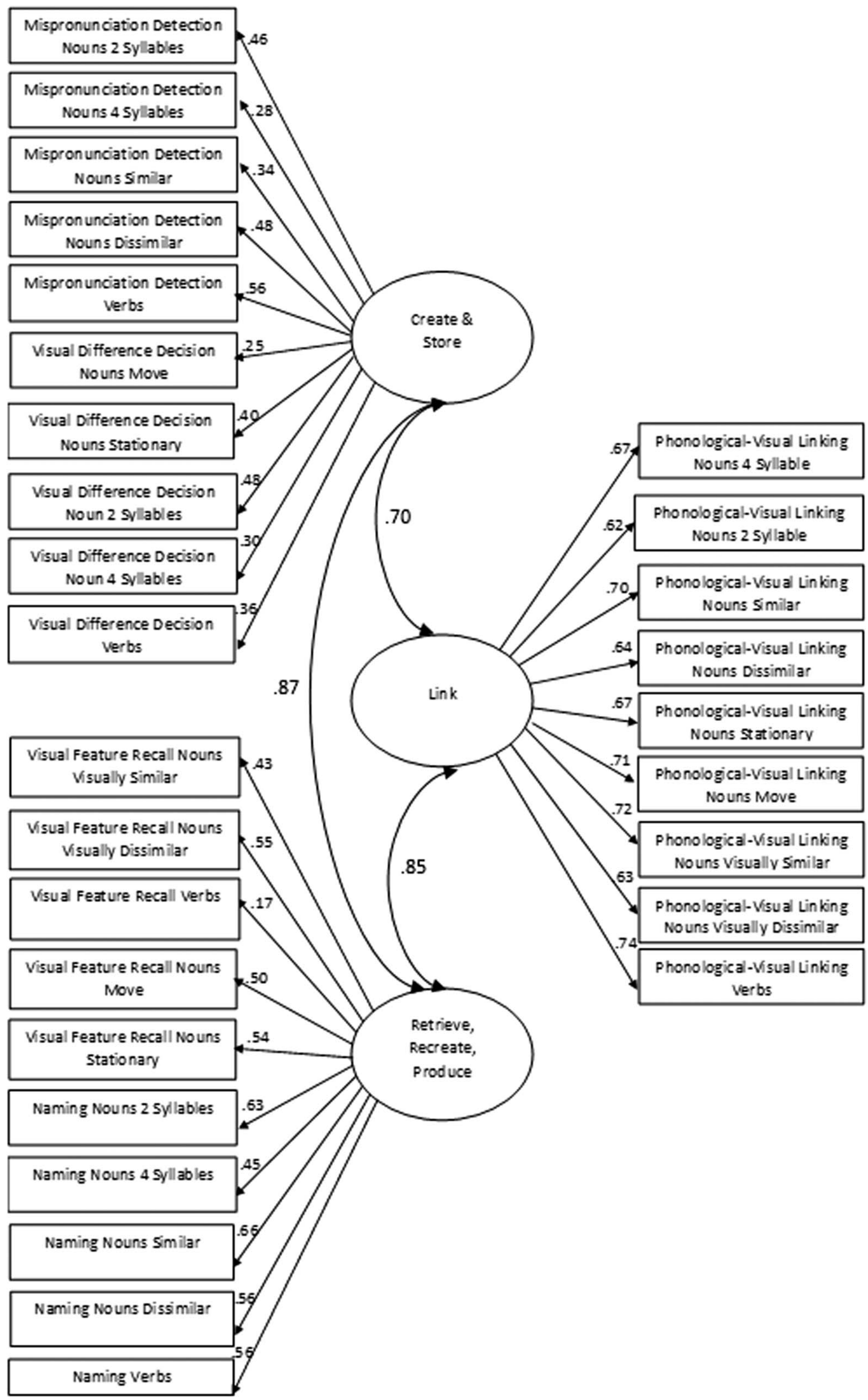


Table 4. Descriptive statistics for variables in the word learning models.

Variable	*M*	*SD*	*n*
Mispronunciation Detections Nouns 2 Syllables	.539	.338	159
Mispronunciation Detections Nouns 4 Syllables	.443	.349	159
Mispronunciation Detection Nouns Phonologically Similar	.341	.340	162
Mispronunciation Detection Nouns Phonologically Dissimilar	.306	.345	162
Mispronunciation Detection Verbs	.387	.306	155
Naming Nouns 2 Syllables	.373	.189	155
Naming Nouns 4 Syllables	.194	.115	155
Naming Nouns Phonologically Similar	.286	.178	158
Naming Nouns Phonologically Dissimilar	.276	.173	157
Naming Verbs	.198	.117	150
Visual Difference Decision Nouns Stationary	.605	.318	160
Visual Difference Decision Nouns Move	.631	.282	160
Visual Difference Decision Nouns Visually Similar	.690	.320	159
Visual Difference Decision Nouns Visually Dissimilar	.675	.299	159
Visual Difference Decision Verbs	.520	.257	155
Visual Feature Recall Nouns Stationary	.691	.165	160
Visual Feature Recall Nouns Move	.681	.175	160
Visual Feature Recall Verbs	.427	.143	155
Phonological–Visual Linking Nouns 2 Syllable	.675	.168	159
Phonological–Visual Linking Nouns 4 Syllable	.714	.148	159
Phonological–Visual Linking Nouns Phonologically Similar	.605	.196	162
Phonological–Visual Linking Nouns Phonologically Dissimilar	.618	.187	162
Phonological–Visual Linking Nouns Visually Similar	.612	.188	162
Phonological–Visual Linking Nouns Visually Dissimilar	.640	.190	162
Phonological–Visual Linking Verbs Stationary	.632	.181	160
Phonological–Visual Linking Verbs Move	.628	.189	160

Step 2: Model Refinement

Refinement was explored for the phonological–semantic latent variable model (Model 3) as this was the best fitting model based on AIC, RMSEA, SRMR, TLI, and the S-B χ^2 difference test. Based on the standardized loadings, we restricted the phonological–visual linking tasks to the phonological factor only. This refined version of the phonological–semantic model (see Figure 7) had fit indices similar to the original version of Model 3 (TLI = 0.919, RMSEA = 0.045 [0.034, 0.055], p = .352, AIC = −2842, SRMR = 0.066). The S-B χ^2 was significant, S-B $\chi^2(9)$ = 19.15, p = .0239, but the refined model explained more variance than the original model (see Table 7). The standardized loadings improved such that all paths were significant and no path loading was greater than 1. Previously reported nonsignificant correlated errors remained nonsignificant with the exception of Mispronunciation Detection Nouns–Phonologically Similar and Dissimilar (B = 0.17, z = 1.98, p = .048). Additionally, the covariance between the phonological and semantic factors reduced to 0.664 (z = 8.67, p < .001), which meant that the two factors were more distinct when the cross loadings were removed. The version that restricted the phonological–visual linking tasks to the phonological factor was thus determined to be the "best" latent model as this version had acceptable model fit and was more parsimonious than the original version.

Discussion

Our word learning games challenged second graders with TD to trigger the word learning process when they heard a new word, to create and store phonological and semantic representations of the word, and to link those representations in memory. We assessed noun and verb learning receptively and expressively. Four key findings about the structure of word learning during triggering and elaboration emerged.

First, although recent research has shown that oral language appears to be unidimensional in nature from Grades 1 to 3 (Anthony et al., 2014; Bornstein et al., 2014; Foorman et al., 2015; LARRC, 2015, 2017; Lonigan & Milburn, 2017; Tomblin & Zhang, 2006), this was not the case for our dynamic measures of word learning. Static measures of language, such as receptive and expressive vocabulary tests, assess the product of learning rather than the learning process, which may be why language scores on multiple measures of language in young children with TD are highly correlated. However, dynamic learning measures permit a more fine-grained assessment of factors contributing to learning. When we designed tasks specifically to test underlying word learning processes, we found that the structure of word learning was not unidimensional.

Second, we did not find evidence for structural differences in receptive and expressive word learning. Rather, Model 2 (see Figure 5) shows that the receptive and expressive word learning factors were highly correlated (.89), indicating that receptive and expressive indicators tapped the same underlying construct. If this is the case, why do many children score higher on receptive than expressive word learning and vocabulary measures? Our results suggest that differences are due to assessment requirements in relation to the robustness of each word's underlying phonological

Table 5. Intercorrelations among working memory variables.

V	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	—	**.391**	.202	**.039**	.104	**.168**	.292	**.230**	**.259**	**.238**	**.115**	–.026	.215	**.195**	**.130**	.078	.291	.242	.269	.323	.182	.188	.233	.170	.291	.267
2	**.000**	—	**.216**	.105	**.056**	**.246**	**.160**	.195	**.069**	**.056**	**–.033**	**.003**	**.159**	.144	**.064**	**–.003**	.160	.227	.317	.253	.038	.064	.224	.134	.143	.089
3	**.012**	**.007**	—	**.381**	.102	**.309**	**.482**	**.435**	**.215**	**.340**	**.262**	**.063**	**.242**	**.290**	.101	**.066**	.232	.225	.452	.446	.372	.370	.397	.446	.465	.362
4	.629	.192	**.000**	—	.147	.274	**.332**	.338	.257	.230	.177	**.167**	.247	.253	.152	.231	.050	.126	.314	.145	.267	.209	.196	.338	.167	.120
5	.194	.485	.212	.070	—	**.236**	.198	**.174**	**.291**	**.162**	**.046**	**.121**	**.175**	**.197**	**.081**	.218	–.003	.319	.144	.101	.268	.234	.160	.189	.237	.140
6	**.037**	**.002**	**.000**	**.001**	**.002**	—	**.292**	.244	**.296**	**.145**	**.150**	**.150**	**.223**	**.182**	**.127**	**.215**	.054	.276	.153	.159	.251	.276	.150	.192	.152	.125
7	**.000**	**.048**	**.000**	**.000**	**.013**	**.000**	—	**.594**	**.249**	**.415**	**.119**	.094	**.320**	**.264**	**.163**	**.144**	.089	.295	.303	.367	.500	.396	.261	.308	.348	.251
8	**.004**	**.016**	**.000**	**.000**	**.029**	**.002**	**.000**	—	**.211**	**.268**	**.138**	**.007**	**.201**	**.135**	**.144**	**.071**	.091	.106	.285	.416	.515	.460	.375	.363	.358	.341
9	**.001**	.403	**.010**	**.002**	**.000**	**.000**	**.002**	**.010**	—	**.340**	**.236**	**.130**	.236	**.149**	**.355**	**.157**	.114	.266	.298	.210	.255	.216	.317	.340	.331	.292
10	**.004**	.501	**.000**	**.006**	.051	.080	**.000**	**.001**	**.000**	—	**.161**	**.107**	**.215**	.290	.248	.215	.247	.184	.316	.203	.222	.173	.317	.364	.371	.233
11	.154	.681	**.001**	**.030**	.563	.061	.140	.088	**.004**	.052	—	.086	.265	.342	.137	.090	.193	.236	.170	.180	.124	.078	.149	.256	.154	.130
12	.745	.975	.441	**.040**	.129	.060	.246	.930	.112	.200	.277	—	.181	.171	.084	.137	.088	.111	.099	.002	.078	.104	.177	.096	.030	.086
13	**.007**	**.048**	**.003**	**.002**	**.028**	**.005**	**.000**	**.013**	**.004**	**.009**	**.001**	**.022**	—	**.438**	.199	.182	.101	.321	.230	.222	.250	.189	.341	.433	.248	.194
14	**.015**	.073	**.000**	**.002**	**.013**	**.022**	**.001**	.096	.069	**.000**	**.000**	**.030**	**.000**	—	**.239**	.195	.194	.364	.272	.260	.199	.124	.266	.356	.265	.224
15	.114	.436	.228	.068	.321	.119	**.047**	.082	**.000**	**.002**	.093	.305	**.014**	**.003**	—	.104	.079	.060	.016	.011	.127	.033	.242	.232	.138	.088
16	.347	.972	.431	**.005**	**.007**	**.008**	.081	.390	.051	**.008**	.269	.094	**.026**	**.016**	.197	—	–.137	.196	.010	.000	.044	–.032	–.054	.036	.111	.064
17	**.000**	**.044**	**.004**	.537	.975	.507	.274	.264	.166	**.003**	**.016**	.278	.209	**.016**	.337	.096	—	.116	.250	.241	.140	.164	.171	.169	.157	.121
18	**.002**	**.004**	**.005**	.119	**.000**	**.000**	**.000**	.193	**.001**	**.027**	**.003**	.168	**.000**	**.000**	.470	**.017**	.145	—	.200	.257	.155	.162	.106	.164	.107	.076
19	**.001**	**.000**	**.000**	**.000**	.074	.056	**.000**	**.000**	**.000**	**.000**	**.034**	.221	**.004**	**.001**	.843	.908	**.002**	**.012**	—	.682	.424	.432	.428	.438	.378	.362
20	**.000**	**.001**	**.000**	.071	.212	**.047**	**.000**	**.000**	**.010**	**.015**	**.025**	.976	**.005**	**.001**	.898	.998	**.002**	**.001**	**.000**	—	.377	.406	.354	.360	.365	.329
21	**.023**	.634	**.000**	**.001**	**.001**	**.001**	**.000**	**.000**	**.001**	**.007**	.120	.328	**.002**	**.012**	.120	.594	.082	.053	**.000**	**.000**	—	.773	.451	.470	.518	.472
22	**.019**	.428	**.000**	**.010**	**.003**	**.000**	**.000**	**.000**	**.008**	**.036**	.329	.194	**.018**	.121	.685	.695	**.041**	**.043**	**.000**	**.000**	**.000**	—	.365	.360	.405	.397
23	**.003**	**.005**	**.000**	**.016**	**.044**	.060	**.001**	**.000**	**.000**	**.000**	.060	**.025**	**.000**	**.001**	**.003**	.512	**.034**	.190	**.000**	**.000**	**.000**	**.000**	—	.778	.453	.416
24	**.034**	.097	**.000**	**.000**	**.017**	**.016**	**.000**	**.000**	**.000**	**.000**	**.001**	.227	**.000**	**.000**	**.004**	.658	**.035**	**.041**	**.000**	**.000**	**.000**	**.000**	**.000**	—	.487	.422
25	**.000**	.074	**.000**	**.039**	**.003**	.056	**.000**	**.000**	**.000**	**.000**	.054	.704	**.002**	**.001**	.090	.172	.051	.186	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	—	.783
26	**.001**	.267	**.000**	.141	.079	.115	**.002**	**.000**	**.000**	**.005**	.103	.282	**.014**	**.005**	.283	.436	.133	.347	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	—

Note. *R* values are reported in the upper triangle and *p* values in the lower triangle. Bolded values indicate $p < .05$. V = variable. Variable names are as follows:

1. Mispronunciation Detection Nouns 2 Syllables
2. Mispronunciation Detections Nouns 4 Syllables
3. Naming Nouns 2 Syllables
4. Naming Nouns 4 Syllables
5. Mispronunciation Detection Nouns Similar
6. Mispronunciation Detection Nouns Dissimilar
7. Naming Nouns Phonologically Similar
8. Naming Nouns Phonologically Dissimilar
9. Mispronunciation Detection Verbs
10. Naming Verbs
11. Visual Differences Decision Nouns Stationary
12. Visual Difference Decision Nouns Move
13. Visual Feature Recall Nouns Stationary
14. Visual Difference Recall Nouns Move
15. Visual Difference Decision Verbs
16. Visual Feature Recall Verbs
17. Visual Difference Decision Nouns Visually Similar
18. Visual Difference Decision Nouns Visually Dissimilar
19. Phonological–Visual Linking Nouns 2 Syllable
20. Phonological–Visual Linking Nouns 4 Syllable
21. Phonological–Visual Linking Nouns Phonologically Similar
22. Phonological–Visual Linking Nouns Phonologically Dissimilar
23. Phonological–Visual Verbs Stationary
24. Phonological–Visual Linking Verbs Move
25. Phonological–Visual Linking Nouns Visually Similar
26. Phonological–Visual Linking Nouns Visually Dissimilar

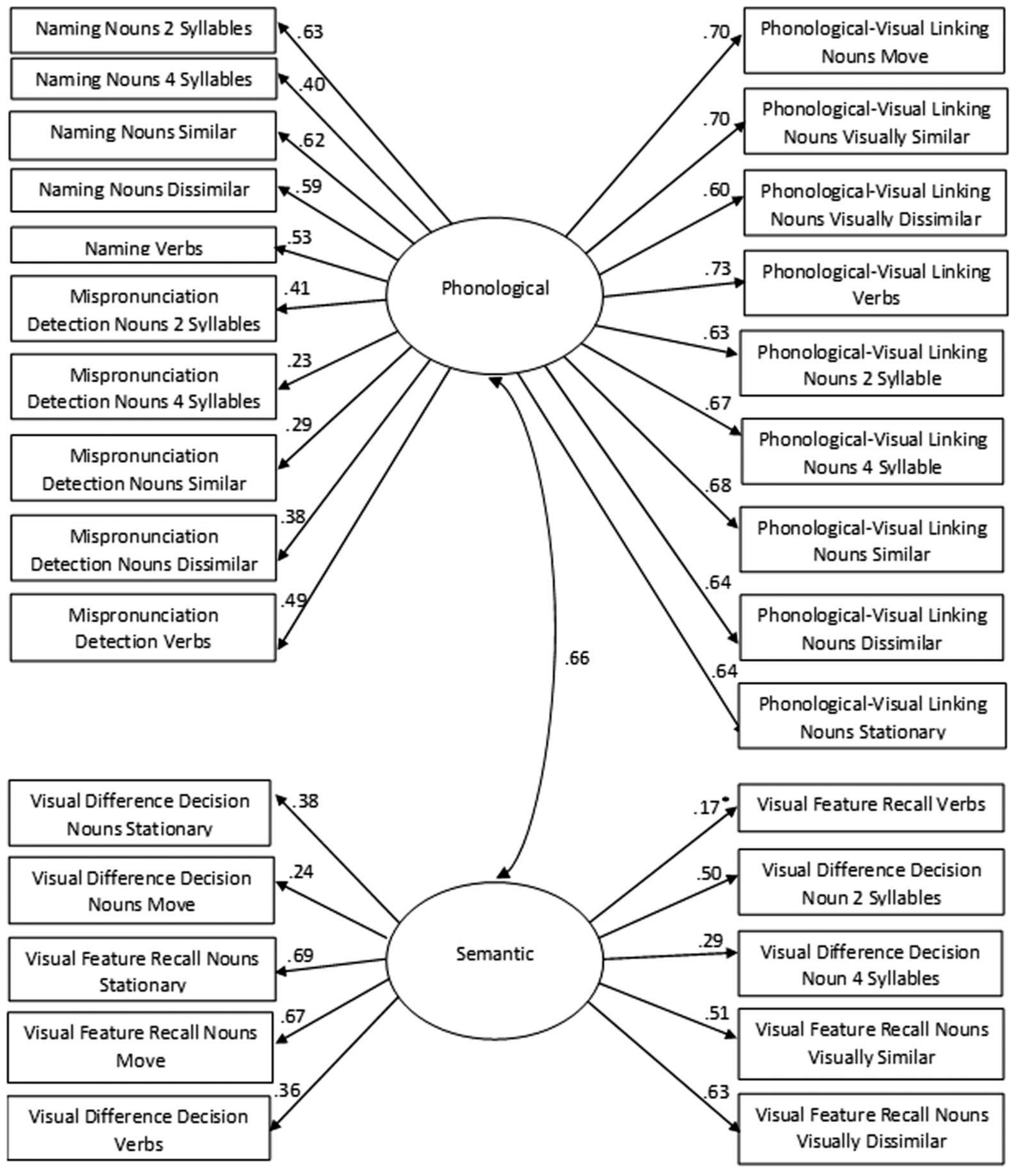
Table 6. Latent factor specification for each variable in the a priori models.

Variable	Receptive & expressive (Model 2)	Phonological & semantic (Model 3)	Create, recreate, & link (Model 4)
Mispronunciation Detections Nouns 2 Syllables	Receptive	Phonological	Create & store
Mispronunciation Detections Nouns 4 Syllables	Receptive	Phonological	Create & store
Naming Nouns 2 Syllables	Expressive	Phonological	Retrieve & produce
Naming Nouns 4 Syllables	Expressive	Phonological	Retrieve & produce
Mispronunciation Detection Nouns Phonologically Similar	Receptive	Phonological	Create & store
Mispronunciation Detection Nouns Phonologically Dissimilar	Receptive	Phonological	Create & store
Naming Nouns Similar	Expressive	Phonological	Retrieve & produce
Naming Nouns Dissimilar	Expressive	Phonological	Retrieve & produce
Mispronunciation Detection Verbs	Receptive	Phonological	Create & store
Naming Verbs	Expressive	Phonological	Retrieve & produce
Visual Difference Decision Nouns Stationary	Receptive	Semantic	Create & store
Visual Difference Decision Nouns Move	Receptive	Semantic	Create & store
Visual Feature Recall Nouns Stationary	Receptive	Semantic	Retrieve & produce
Visual Feature Recall Nouns Move	Receptive	Semantic	Retrieve & produce
Visual Difference Decision Verbs	Receptive	Semantic	Create & store
Visual Feature Recall Verbs	Receptive	Semantic	Retrieve & produce
Visual Difference Decision Nouns Visually Similar	Receptive	Semantic	Create & store
Visual Difference Decision Nouns Visually Dissimilar	Receptive	Semantic	Create & store
Visual Feature Recall Nouns Visually Similar	Receptive	Semantic	Retrieve & produce
Visual Feature Recall Nouns Visually Dissimilar	Receptive	Semantic	Retrieve & produce
Phonological–Visual Linking Nouns 2 Syllable	Receptive	Phonological, semantic	Link
Phonological–Visual Linking Nouns 4 Syllable	Receptive	Phonological, semantic	Link
Phonological–Visual Linking Nouns Phonologically Similar	Receptive	Phonological, semantic	Link
Phonological–Visual Linking Nouns Phonologically Dissimilar	Receptive	Phonological, semantic	Link
Phonological–Visual Linking Verbs Stationary	Receptive	Phonological, semantic	Link
Phonological–Visual Linking Verbs Move	Receptive	Phonological, semantic	Link
Phonological–Visual Linking Nouns Visually Similar	Receptive	Phonological, semantic	Link
Phonological–Visual Linking Nouns Visually Dissimilar	Receptive	Phonological, semantic	Link

and semantic representations. For receptive measures, when a child is asked to point to the referent for a word, the phonology of the word is provided, as is the semantic representation. This means that the child does not have to recall or produce either representation but only activate their link. This makes it possible to respond correctly with weaker representations. For expressive measures, however, when shown a referent and asked to name it, the child must recall the phonological representation, recreate the word, and produce the word's phonology with no support. A correct answer relies on the recall of a detailed, phoneme-by-phoneme representation of the word. In our study, we also assessed whether children could recall, recreate, and produce the semantic representation of a word when given the name (phonology). Here, a correct response relies on the recall of detailed visual and spatial information about the referent. Thus, differences in performance on receptive and expressive word learning and vocabulary measures are accounted for by task demands rather than different underlying constructs. This interpretation is consistent with Marulis and Neuman's (2010) meta-analysis of vocabulary interventions that did not find differences in treatment effects for receptive versus expressive measures. It is also consistent with Leonard's (2009) proposal that separate receptive and expressive language disorders are not likely to exist. It follows that, when teaching a new word, parents or educators are not helping children build separate receptive and expressive representations. Rather, they are strengthening the phonological and semantic representations of each word and the link between them.

Third, we hypothesized that, because word learning requires both phonological and semantic skills and because these skills are associated with different processing areas in the brain, Model 3 with differentiated phonological and semantic factors might best fit the data. This was contrasted with Model 4 that tested whether factors related to the word learning process, including the creation and storage of new phonological and semantic representations, the linking of those representations, and the recreation, retrieval, and production of new nonwords, would best fit the data. By considering the word learning process, Model 4 is consistent with the idea of functional learning networks in the brain. Pulvermuller (1999) proposed that, when neurons in different cortical areas are repeatedly activated at the same time, as occurs when processing a word, connected cell assemblies form and become a functioning unit. This is known as Hebbian learning (Hebb, 1949). Once a cell assembly has formed for a word, it may be activated by incoming phonological or semantic information that spreads activation throughout the network. The fit of Model 4 suggests that

Figure 7. Refined version of the phonological–semantic model (Model 3 refined). *Indicates nonsignificant path loading.



this is a plausible representation of the triggering and elaboration stage of word learning; however, Model 3 better fits the data with lower SRMR, RMSEA, and AIC values. This may indicate that, at the earlier stages of word learning, the strength of the phonological and semantic representations, which are necessary to activate the correct cell assemblies for new words, better represent the word learning process. A testable hypothesis is whether Model 4 would best fit the data during the engagement stage of word learning when phonological and semantic representations are sufficiently established to activate phonologically or semantically related words in the lexicon.

Fourth, we found that restricting phonological–visual indicator loadings to the phonological factor improved the fit of Model 3 by making the phonological and semantic factors more distinct. This finding, together with the high loadings of linking task indicators on the phonological factor (.64–.73), suggests that, at the early stages of word learning, successful linking may depend more heavily on phonological than semantic representations.

Several other interesting observations relate to the strength of item loadings on the phonological and semantic factors in refined Model 3. The first is that, consistent with prior research (e.g., Bornstein et al., 2004; Childers & Tomasello, 2006; Gentner, 1982, 2006; Maguire et al., 2005), nouns appeared to be easier to learn than verbs, probably due in part to the higher cognitive load conveyed when learners must attend to both an action and the label

Table 7. Goodness of fit statistics and model comparisons.

Model	χ^2	TLI	AIC	RMSEA	SRMR	AVE	S-B χ^2
Model 1	540.67			0.053	0.071	0.201	
Base–unidimensional	(*df* = 365)***	0.887	−2794.12	[0.044, 0.063]			
Model 2	535.88			0.053	0.070	0.206	3.89
Receptive–expressive	(*df* = 364)***	0.889	−2797.41	[0.042, 0.062]			(*df* = 1)*
Model 3 (original)	469.57			0.044	0.062	0.234	61.01
Phonological–semantic	(*df* = 355)***	0.925	−2847.25	[0.032, 0.054]			(*df* = 10)***
Model 3[a] (refined)	490.45	0.919	−2841.77	0.045	0.066	0.228	19.15
Phonological–semantic	(*df* = 364)***			[0.034, 0.055]			(*df* = 9)*
Model 4	503.66			0.048	0.067	0.242	31.19
Create–link–retrieve	(*df* = 362)***	0.908	−2825.64	[0.038, 0.058]			(*df* = 3)***

Note. All values reported are based on robust statistics for nonnormality and using full information maximum likelihood to accommodate missing data. TLI = Tucker–Lewis Index; AIC = Akaike information criteria; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean residual; AVE = average variance extracted; S-B χ^2 = Satorra–Bentler rescaled χ^2.

[a]Compared model fit of refined Model 3 to original Model 3; all other comparisons were compared to Model 1. *$p < .05$. ***$p < .001$.

for that action (Childers & Tomasello, 2006). In our study, mean scores for verb naming, visual difference decision for verbs, and visual feature recall for verbs in Table 4 were on the low end compared to nouns. Paired with this, we found that phonological–visual linking for verbs had a higher loading (.73) on the phonological factor than any other linking indicator and that mispronunciation detection for verbs (.49) had the highest loading on the phonological factor of any mispronunciation task. Verb indicator loadings on the semantic factor did not stand out. This suggests that, during the triggering and configurations stages of word learning, successful verb learning relies more heavily on the creation and storage of phonological representations than semantic representations. Said another way, it may be more difficult to create, store, and recall the phonological label for the action than to create, store, and recall the action itself.

A second interesting observation was that, in general, naming task indicators had higher loadings on the phonological factor (.40–.62) than mispronunciation detection tasks (.23–.49) that required recognition of the correct name. Because naming tasks are a more stringent test of the specificity and strength of phonological representations than mispronunciation tasks, it is not surprising that they were more strongly correlated with the underlying phonological factor. Similarly (except for verbs), visual feature recall task indicators that required children to recreate the referent had higher loadings on the semantic factor (.51–.69) than visual difference decision indicators (.24–.50) that required recognition of the referent. We designed the visual feature recall task for semantic representations to be analogous to naming for phonological representations. Thus, tasks requiring children to produce a representation were more highly correlated with the underlying factor than those requiring recognition.

Limitations

Finally, we must acknowledge limitations in what a structural equation modeling approach can do. First, when we build a statistical model, we are not actually saying "this is how the world is." Instead, we are saying "this is how we are going to reason about the world, in light of our data, theory, constraints, and purposes." It is always the case that there are more complex phenomena in the real world than what we build into our model.

Second, structural equation modeling examines processes that differ among individuals. For the models we have rejected, we avoid the strong claim that they reflect nonexistent processes. Clearly, the process of learning words is likely to involve receptive understanding, associative linking, memory storage, retrieval, and expressive processes. Instead, our more limited claim, based on our analyses and results, is that an important method of examining "individual differences" in word learning ability is through the dimensions of phonological and semantic processing.

Third, the labels we have given to the factors in our "winning" model may well be oversimplifications of what those factors reflect. The phonological factor includes not only the ability to remember phonological information per se but also the ability to link together phonological and nonphonological representations as is needed for the naming and linking tasks. Indeed, the tasks that require linking uniformly load more highly on this factor than the tasks that focus on mispronunciation detection. It will require further work to determine whether pure phonological memory and phonological–semantic linking are perfectly related or are separate subfactors within word learning performance. Moreover, the semantic factor relies solely on semantic differences as they are reflected visually, and it will require further work to determine whether a nonvisual (e.g., verbal) test of the semantic features would yield similar results.

Finally, our results inform the structure of word learning in second-grade children with TD, but over the course of development or in groups of children with developmental disorders, it is possible that the structure of word learning varies. We are in the process of testing this in our ongoing studies.

Conclusions

The structure of word learning during triggering and configuration in second-grade children with TD reflects separate but related phonological and semantic factors. We did not find evidence for a unidimensional latent variable model of word learning or for separate receptive and expressive word learning factors. In future studies, it will be interesting to determine whether the structure of word learning differs during the engagement stage of word learning when phonological and semantic representations, as well as the links between them, are sufficiently strong to affect other words in the lexicon.

Acknowledgments

This work was funded by a grant from the National Institutes of Health - National Institute of Deafness and Other Communication Disorders to Principal Investigator Shelley Gray (R01 DC010784). We sincerely thank the children and their families who participated in this research and the following school districts and educational organizations who partnered with us to recruit participants: Academy of Math and Science, Alhambra School District, Arizona Museum for Youth, Chandler Unified School District, Desert Springs Academy, Foothills Academy, City of Tempe Kid's Zone Enrichment Program, Concordia East Valley Children's Theatre Boy's and Girl's Clubs, Gilbert Public Schools, Glendale Elementary School District, Imagine Schools, Isaac School District, KIDCO, La Paloma Academy, Maricopa Unified School District, Mesa Public Schools, Nogales Unified School District, San Tan Charter School, Satori, Sunnyside Unified School District, Tempe School District, Tucson Unified School District, Vail Unified School District, and Wildcat School.

References

Abel, A. D., Schneider, J., & Maguire, M. J. (2017). N400 response indexes word learning from linguistic context in children. *Language Learning and Development, 14*(1), 61–71. https://doi.org/10.1080/15475441.2017.1362347

Alt, M., Arizmendi, G. D., Gray, S., Hogan, T. P., Green, S., & Cowan, N. (2019). Novel word learning in children who are bilingual: Comparison to monolingual peers. *Journal of Speech, Language, and Hearing Research, 62*(7), 2332–2360. https://doi.org/10.1044/2019_JSLHR-L-18-0009

Alt, M., Gray, S., Hogan, T. P., & Nelson, C. (2019). Spoken word learning differences among children with dyslexia, concomitant dyslexia and developmental language disorder, and typical development. *Language, Speech, and Hearing Services in Schools, 50*(4), 540–561. https://doi.org/10.1044/2019_LSHSS-VOIA-18-0138

Alt, M., Hogan, T., Green, S., Gray, S., Cabbage, K., & Cowan, N. (2017). Word learning deficits in children with dyslexia. *Journal of Speech, Language, and Hearing Research, 60*(4), 1012–1028. https://doi.org/10.1044/2016_JSLHR-L-16-0036

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.).

Anthony, J. L., Davis, C., Williams, J. M., & Anthony, T. I. (2014). Preschoolers' oral language abilities: A multilevel examination of dimensionality. *Learning and Individual Differences, 35,* 56–61. https://doi.org/10.1016/j.lindif.2014.07.004

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory. *Trends in Cognitive Sciences, 4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1.

Baron, L. S., Hogan, T. P., Alt, M., Gray, S., Cabbage, K. L., Green, S., & Cowan, N. (2018). Children with dyslexia benefit from orthographic facilitation during spoken word learning. *Journal of Speech, Language, and Hearing Research, 61*(18), 2002–2014. https://doi.org/10.1044/2018_JSLHR-L-17-0336

Batterink, L., & Neville, H. (2011). Implicit and explicit mechanisms of word learning in a narrative context: An event-related potential study. *Journal of Cognitive Neuropsychology, 23*(11), 3181–3196. https://doi.org/10.1162/jocn_a_00013

Bloom, P. (2000). *How children learn the meanings of words*. MIT Press.

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. R., & Mesulam, M. M. (2002a). Functional anatomy of intra- and cross-modal lexical tasks. *NeuroImage, 16*(1), 7–22. https://doi.org/10.1006/nimg.2002.1081

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. R., & Mesulam, M. M. (2002b). Modality independence of word comprehension. *Human Brain Mapping, 16*(4), 251–261. https://doi.org/10.1002/hbm.10054

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. R., & Mesulam, M. M. (2003). Relation between brain activation and lexical performance. *Human Brain Mapping, 19*(3), 155–169. https://doi.org/10.1002/hbm.10111

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2004). Development of brain mechanisms for processing orthographic and phonologic representations. *Journal of Cognitive Neuroscience, 16*(7), 1234–1249. https://doi.org/10.1162/0898929041920496

Booth, J. R., Burman, D. D., Meyer, J. R., Zhang, L., Choy, J., Gitelman, D. R., Parrish, T. R., & Mesulam, M. M. (2003). Modality-specific and -independent developmental differences in the neural substrate for lexical processing. *Journal of Neurolinguistics, 16*(4–5), 383–405. https://doi.org/10.1016/S0911-6044(03)00019-8

Born, J., Rasch, B., & Gais, S. (2006). Sleep to remember. *Neuroscientist, 12,* 410–424. https://doi.org/10.1177/1073858406292647

Bornstein, M., Cole, L., Maital, S. K., Park, S. Y., Pascual, L., Pêcheux, M. G., Ruel, J., Venuti, P., & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean and American English. *Child Development, 75*(4), 1115–1139. https://doi.org/10.1111/j.1467-8624.2004.00729.x

Bornstein, M. H., Hahn, C., Putnick, D. L., & Suwalsky, J. T. D. (2014). Stability of core language skill from early childhood to adolescence: A latent variable approach. *Child Development, 85*(4), 1346–1356. https://doi.org/10.1111/cdev.12192

Breitenstein, C., Jansen, A., Deppe, M., Foerster, A. F., Sommer, J., Wolbers, T., & Knecht, S. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage, 25*(3), 958–568. https://doi.org/10.1016/j.neuroimage.2004.12.019

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). SAGE.

Byrne, B. M. (2012). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. Springer.

Cabbage, K. L., Brinkley, S., Gray, S., Alt, M., Cowan, N., Green, S., Kuo, T., & Hogan, T. (2017). Assessing working memory in children: The Comprehensive Assessment Battery for Children–Working Memory (CABC-WM). *Journal of Visualized Experiments, 124,* e55121. https://doi.org/10.3791/55121

Carey, S. (2010). Beyond fast mapping. *Language Learning and Development, 6*(3), 184–205. https://doi.org/10.1080/15475441.2010.484379

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference*, *15*, 17–29.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Childers, J. B., & Tomasello, M. (2006). Are nouns easier to learn than verbs? Three experimental studies. In K. A. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 1–36). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195170009.003.0013

Cowan, N. (2001). The magical number four in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87–114. https://doi.org/10.1017/s0140525x01003922

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1536), 3773–3800. https://doi.org/10.1098/rstb.2009.0111

DuPaul, G. J., Anastopoulos, A. D., Power, T. J., Reid, R., Ikeda, M. J., & McGoey, K. E. (1998). Parent ratings of attention deficit/hyperactivity disorder symptoms: Factor structure and normative data. *Journal of Psychopathology and Behavioral Assessment, 20*(1), 83–102.

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD Rating Scale–IV (ADHD-RS)*. Guilford.

Erikson, J. A., Alt, M., Gray, S., Green, S., Hogan, T. P., & Cowan, N. (2018). Phonological vulnerability for school-aged Spanish–English speaking bilingual children. *International Journal of Bilingual Education and Bilingualism*. https://www.tandfonline.com/doi/full/10.1080/13670050.2018.1510892

Fazly, A., Alishahi, A., & Stevenson, S. (2010). A probabilistic computational model of cross-situational word learning. *Cognitive Science, 34*(6), 1017–1063. https://doi.org/10.1111/j.1551-6709.2010.01104.x

Foorman, B. R., Herrera, S., Petscher, Y., Mitchell, A., & Truckenmiller, A. (2015). The structure of oral language and reading and their relation to comprehension in kindergarten through grade 2. *Reading and Writing: An Interdisciplinary Journal, 28,* 655–681. https://doi.org/10.1007/s11145-015-9544-5

Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science, 20*(5), 578–585. https://doi.org/10.1111/j.1467-9280.2009.02335.x

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*(2), 105–132. https://doi.org/10.1016/S0010-0277(03)00070-2

Gentner, D. (1982). Why nouns are learned before verbs: Linguistical relativity versus natural partitioning. In S. A. Kuczaj, II (Ed.), *Language development: Vol. 2. Language, thought, and culture* (pp. 301–334). Erlbaum.

Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek & R. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 544–564). Oxford University Press.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation–Second Edition (GFTA-2)*. AGS. https://doi.org/10.1037/t15098-000

Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6–10. https://doi.org/10.1177/074193258600700104

Gray, S., Alt, M., Hogan, T. P., Green, S., & Cowan, N. (2020). *Comprehensive Assessment Battery for Children–Working memory* [Manuscript in preparation]. Arizona State University, Tempe.

Gray, S., Pittman, A., & Weinhold, J. (2014). Effect of phonotactic probability and neighborhood density on word-learning by preschoolers with typical development and specific language impairment. *Journal of Speech, Language, and Hearing Research, 57*(3), 1011–1025. https://doi.org/10.1044/2014_JSLHR-L-12-0282

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Pearson.

Hebb, D. O. (1949). *The organization of behavior*. Wiley.

Hollich, G. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model of the origins of word learning. *Monographs of the Society for Research in Child Development, 65*(3), 17–29. https://doi.org/10.1111/1540-5834.00092

Hoover, W., & Gough, P. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2,* 127–160. https://doi.org/10.1007/BF00401799

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*(1), 100–116. https://doi.org/10.1016/j.jml.2010.02.003

Hoyle, R. H. (1995). *Structural equation modelling—Concepts, issues, and applications*. SAGE.

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the sematic maps that tile human cerebral cortex. *Nature, 532,* 453–458. https://doi.org/10.1038/nature17637

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology, 2,* 1–16. https://doi.org/10.3389/fpsyg.2011.00255

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*(3), 739–756. https://doi.org/10.1044/1092-4388(2009/08-0248)

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children, Second Edition (K-ABC2)*. AGS.

Language and Reading Research Consortium. (2015). The dimensionality of language ability in young children. *Child Development, 86*(6), 1948–1965. https://doi.org/10.1111/cdev.12450

Language and Reading Research Consortium. (2017). Oral language and listening comprehension: Same or different constructs. *Journal of Speech, Language, and Hearing Research, 60*(5), 1273–1284. https://doi.org/10.1044/2017_JSLHR-L-16-0039

Leach, L., & Samuel, A. G. (2007). Lexical configuration and lexical engagement: When adults learn new words. *Cognitive Psychology, 55*(4), 306–353. https://doi.org/10.1016/j.cogpsych.2007.01.001

Leonard, L. B. (2009). Is expressive language disorder an accurate diagnostic category. *American Journal of Speech-Language Pathology, 18*(2), 115–123. https://doi.org/10.1044/1058-0360(2008/08-0064)

Lin, Y. (2015). The acquisition of words' meaning based on constructivism. *Theory and Practice in Language Studies, 5*(3), 639–645. https://doi.org/10.17507/tpls.0503.26

Lonigan, C. J., & Milburn, T. F. (2017). Identifying the dimensionality of oral language skills of children with typical development in preschool through fifth grade. *Journal of Speech, Language, and Hearing Research, 60*(8), 2185–2198. https://doi.org/10.1044/2017_JSLHR-L-15-0402

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51,* 201–226. https://doi.org/10.1146/annurev.psych.51.1.201

Magro, L. O., Attout, L., Majerus, S., & Szmalec, A. (2018). Short- and long-term memory determinants of novel word form learning. *Cognitive Development, 47,* 146–157. https://doi.org/10.1016/j.cogdev.2018.06.002

Maguire, M. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2005). A unified theory of word learning: Putting verb acquisition in context. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. Oxford University Press.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press.

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research, 80*(3), 300–335. https://doi.org/10.3102/0034654310377087

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

Merriman, W. E., & Bowman, L. (1989). The mutual exclusivity bias in children's early word learning. *Monographs of the Society for Research in Child Development, 54*(3–4), 1–132. https://doi.org/10.2307/1166130

Mestres-Misse, A., Rodriquez-Fornells, A., & Munte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex, 17*(8), 1858–1866. https://doi.org/10.1093/cercor/bhl094

Norris, D., Page, P. A., & Hall, J. (2017). Learning nonwords: The Hebb repetition effect as a model of word learning. *Memory, 26,* 852–857. https://doi.org/10.1080/09658211.2017.1416639

Partanen, E., Leminen, A., de Paoli, S., Buyndgaard, A., Kingo, O. S., Krojgaard, P., & Shtyrov, Y. (2017). Flexible, rapid and automatic neocortical word form acquisition mechanism in children as revealed by neuromagnetic brain response dynamics. *NeuroImage, 155*(15), 450–459. https://doi.org/10.1016/j.neuroimage.2017.03.066

Pulvermuller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences, 22*(2), 253–279. https://doi.org/10.1017/S0140525X9900182X

RStudio Team. (2016). *RStudio: Integrated development environment for R*. RStudio. http://www.rstudio.com/

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. http://www.jstatsoft.org/v48/i02/

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals–Fourth Edition (CELF-4)*. AGS.

Shtyrov, Y. (2011). Fast mapping of novel words forms traced neurophysiologically. *Frontiers in Psychology, 2,* 1–9. https://doi.org/10.3389/fpsyg.2011.00340

Shtyrov, Y. (2012). Neural bases of rapid word learning. *The Neuroscientist, 18,* 312–319. https://doi.org/10.1177/1073858411420299

Shtyrov, Y., Nikulin, V. V., & Pulvermüller, F. (2010). Rapid cortical plasticity underlying novel word learning. *Journal of Neuroscience, 30*(50), 16864–16867. https://doi.org/10.1523/JNEUROSCI.1376-10.2010

Suzuki, W. A. (2006). Encoding new episodes and making them stick. *Neuron, 50*(1), 19–21. https://doi.org/10.1016/j.neuron.2006.03.029

Tarka, P. (2018). An overview of structural equation modeling: Its beginnings, historical development, usefulness and controversies in the social sciences. *Quality & Quantity, 52,* 313–354. https://doi.org/10.1007/s11135-017-0469-8

Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics, 10*(4), 401–413. https://doi.org/10.1075/prag.10.4.01tom

Tomasello, R., Garagnani, M., Wennekers, T., & Pulvermuller, F. (2017). Brain connections of words, perceptions and actions: A neurobiological model of spatio-temporal semantic activation in the human cortex. *Neuropsychologial, 98,* 111–129. https://doi.org/10.1016/j.neuropsychologia.2016.07.004

Tomblin, J. B., & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research, 49,* 1193–1208. https://doi.org/10.1044/1092-4388(2006/086)

Torgesen, J., Wagner, R., & Rashotte, C. (2012). *Test of Word Reading Efficiency–Second Edition*. AGS.

Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. *Child Development, 61*(5), 1461–1473. https://doi.org/10.2307/1130756

Weighall, A. R., Henderson, L. M., Barr, D. J., Cairney, S. A., & Gaskell, M. G. (2017). Eye-tracking the time-course of novel word learning and lexical competition in adults and children. *Brain and Language, 167,* 13–27. https://doi.org/10.1016/j.bandl.2016.07.010

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

Woodcock, R. (2011). *Woodcock Reading Mastery Test–Third Edition*. Pearson.

Yu, C., & Ballard, D. H. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing, 70*(13–15), 2149–2165. https://doi.org/10.1016/j.neucom.2006.01.034