

The Role of Semantic Information in Children's Word Reading:  
Does Meaning Affect Readers' Ability to Say Polysyllabic Words Aloud?

Devin Kearns

Reem Al Ghanem

University of Connecticut

Author Note: Devin M. Kearns, Department of Educational Psychology, University of Connecticut; Reem Al Ghanem, Department of Educational Leadership, University of Connecticut.

This research was made possible by the hard work of research assistants Brian L. Kearns, Michael Li, Jeniffer Cruz, Gabriel Garcia, Natalie Sutton, and Kelsy K. Lua. The authors would also like to acknowledge LaShawna Thompson for conducting supplemental analyses and Donald L. Compton, Jennifer K. Gilbert, and Lesly Wade-Woolley for their ideas for measuring item-specific word knowledge. In addition, the authors would like to thank the administrators who supported this project, the teachers who took time to participate, the parents who permitted their children to participate, and the children who all tried their best.

Correspondence concerning this article should be addressed to Devin M. Kearns, Department of Educational Psychology, Neag School of Education, University of Connecticut, 249 Glenbrook Road, Unit 3064, Storrs, CT 06269.

E-mail: [devin.kearns@uconn.edu](mailto:devin.kearns@uconn.edu)

**ACCEPTED FOR PUBLICATION AUGUST 30, 2018**

## Abstract

In an effort to improve oral reading, beginning and remedial reading programs in English focus on phonological awareness skills and recoding with grapheme-phoneme correspondences. The meanings of the words children practice reading aloud are given little emphasis. Some studies now suggest semantic knowledge may have a direct effect on children's oral reading, but it is unclear whether it is due to knowledge of a given word, general semantic knowledge (vocabulary size), or morphological awareness. We asked third and fourth graders with reading difficulty and their typically-achieving peers ( $N = 95$ ) to read polysyllabic words ( $N = 48$ ) in isolation. We tested children's semantic knowledge for those specific words, general semantic knowledge (vocabulary size), morphological awareness, and orthographic and phonological knowledge. Using generalized linear mixed-effects models, we found a word-specific semantic effect—along with word-specific orthographic and phonological effects—and general effects of semantic knowledge, morphological awareness, and phonological awareness. The results add to the studies showing the importance of semantic information but is unique in clarifying that a general semantic effect may be at least partly morphological. The findings support a distributed processing account of reading acquisition in which readers use all reliable information to pronounce words, not only letter-sound consistencies. There may be implications for curriculum design. The word-specific semantic effect may suggest that beginning readers should practice reading words in their phonological lexicons. The general morphological effect suggests that children might benefit from learning morphological units early in their reading development.

Key words: Word reading; beginning reading instruction; reading difficulty; semantic knowledge

### Educational Impact and Implications Statement

In many English reading programs for beginning readers, children learn phonics, a way of reading words aloud by matching letters and sounds and “sounding out” the words (reading *vat* by connecting the sounds of *v*, *a*, and *t*), but some people think children should spend less time learning phonics and spend more time learning vocabulary. We asked children to read 48 words aloud and if they knew the meanings of those 48 words. We also tested if they knew many vocabulary words, and if they knew how to use meaning parts in words (like how *replacement* is from *place* with *re-*, and *-ment* added).

- We found that children read words better when they knew their meanings, so teachers might want to avoid having children read words that are very rare (like *vat*) even if they are useful for reading practice. They might also want to tell children a little about the meaning of some words before starting a lesson.
- Children also read words better when they understood how to use meaning parts in words, so they might benefit from learning to use meaning parts to read aloud.
- The results do not resolve the debate about how much phonics learning or vocabulary learning children should do, but extended practice linking letters and sounds and linking letters and meaning parts will likely help students read words quickly and accurately.

### The Role of Semantic Information in Children's Word Reading:

#### Does Meaning Affect Readers' Ability to Say Polysyllabic Words Aloud?

For beginning and remedial English reading instruction, educators focus primarily on developing skill in word reading, that is, saying printed words aloud. This emphasis seems wholly appropriate: As Share (1995) made clear, reading acquisition requires many experiences successfully recoding individual words. The English orthographic system is a cipher for its phonology (Gough, Juel, & Griffith, 1992) with “exemplary regularities” (Perfetti, 2003, p. 18). English readers, therefore, acquire orthographic representations of new words primarily by recoding them. This is particularly important for the many low frequency words that collectively comprise the majority of words in English texts even novice readers encounter (Fitzgerald, Elmore, Relyea, Hiebert, & Stenner, 2016). Thus, successful beginning reading requires the acquisition of context-sensitive letter-sound relationships that are used to recode (Perfetti, 1992; Rastle & Coltheart, 1999). Underlying the ability to read using these correspondences is the ability to detect and manipulate spoken words' sublexical phonological units (Anthony et al., 2002; Muter, Hulme, Snowling, & Stevenson, 2004; Perfetti, Beck, Bell, & Hughes, 1987). These phonological skills and knowledge of grapheme-phoneme correspondences are essential for acquiring the ability to read words using the alphabetic principle. On this basis, beginning and remedial reading programs emphasize phonics and phonological awareness skills much more than other aspects of reading (e.g., Gillingham & Stillman, 2014; Wilson, 2005).

The issue in the current paper is that many such programs include a limited emphasis on the semantic aspect of reading. A growing amount of empirical data suggests semantic processes may have associations with word reading skill, beyond the contribution of orthographic and phonological knowledge and skills, a point highlighted by Taylor, Duff, Woollams, Monaghan,

and Ricketts (2015). In this paper, we examine the role of semantic information in reading aloud, with emphasis on the level of semantic information that is involved. Specifically, we consider whether word reading accuracy differences related to meaning are related to a reader's *word-specific semantic knowledge* or their *general semantic knowledge*. The goal is to determine whether one of these types of knowledge, both, or none affects a reader's ability to pronounce a given word. Determining which types of knowledge affect word reading is important because it will have implications for developing better ways to teach children beginning reading skills.

We begin with the idea that reading words is a distributed process that occurs through a system that includes orthographic, phonological, and semantic information about the words, knowledge that varies in quality from word to word, as Perfetti's (2007) lexical quality hypothesis suggests. In this system, these sources of knowledge all connect (e.g., letters represent sounds), and readers improve their word knowledge (build lexical quality) by improving the connections they make between the letters, sounds, and meanings. Readers make these improvements every time they see a word. Reading a word strengthens the reader's word-specific knowledge, such that the given word will be read correctly every time after the first few encounters with it (Manis, 1985). The distributed nature of the system means that reading one word correctly will subtly change the reader's letter, sound, and meaning knowledge related to all words—the reader's general knowledge. The system improves incrementally with extensive practice. Researchers have simulated reading using models with this structure and shown they produce pronunciations that parallel human behavior (e.g., Harm & Seidenberg, 1999, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989).

This conceptualization of reading behavior has important consequences for our

examination of semantic effects in word reading, relative to phonological ones. In English, the orthographic-phonological (OP) connections have a strong effect on the pronunciation the reader gives because the letters represent sounds. By contrast, the orthographic-semantic (OS) connections seem unlikely to affect pronunciation because English letters do not represent concepts in the same way as Chinese characters (McBride-Chang & Kail, 2002) or consonantal roots in Arabic (Al Ghanem & Kearns, 2015). On this basis, it would be possible to argue that the OS part of the distributed system would not affect readers' ability to pronounce written words. However, such an argument against the use of the OS network for pronunciation ignores two characteristics of English orthography, that it is both inconsistent and deep.

### **Value of Semantic Knowledge in an Inconsistent Orthography**

There is little doubt that the OP system is not always trustworthy in English (Waters, Seidenberg, & Bruck, 1984), especially when compared with all other alphabetic European languages (Seymour, Aro, Erskine, & the COST Action A8 Network, 2003). To wit, Rau and colleagues (2015) have shown that English readers are more likely to process words at the lexical level than German readers. Goswami, Porpodas, and Wheelwright (1997) observed a similar phenomenon when comparing reading in English-speaking and Greek-speaking children. One way to facilitate reading in English then might be to use semantic information. Readers could connect letters and sounds by identifying the concepts the letters represent and subsequently linking these concepts to sounds.

Readers almost certainly use semantic information to pronounce written words accurately (Taylor et al., 2015), but it is less clear how readers might exploit semantic knowledge when reading an unknown word. Semantic information could facilitate word pronunciation in two possible ways, either via (1) word-specific semantic information or (2) general semantic

knowledge. Children might produce the correct pronunciation of a given word because they already have some semantic information about the word even if they have never read it, or they might produce the correct pronunciation because they have a strong overall semantic system. Data support both possibilities.

**Word-specific semantic information in word reading.** When a reader encounters an unknown printed word, prior semantic and phonological knowledge of the given word might support accurate pronunciation because the reader has a lexical reference to help resolve any ambiguities in the pronunciation (Kearns, Rogers, Al Ghanem, & Koriakin, 2016). Studies have examined this particular possibility, whether readers' word-specific semantic knowledge affects their ability to pronounce their written forms. Duff and Hulme (2012) had five- and six-year-olds read words varying in OP consistency and imageability. Over six trials, the children were more accurate on consistent words in the first few trials and imageable words in the later ones. The children may have used OP information to recode in the initial trials (Manis, 1985; Share, 1995) but relied on the OS pathway over time because the OS network has been calibrated to permit access to the written word and its associated concepts with little ambiguity.

Wang, Nickels, Nation, and Castles (2013) also found data supporting the idea that letter-sound inconsistency might result in more use of semantic knowledge to support word reading. In their study, Grade 2 children read nonwords with consistent or inconsistent OP mappings. Before seeing the printed words, the children learned the nonwords' pronunciations or meanings. Later, children taught the meanings showed better performance on the inconsistent nonwords but not the consistent ones. Ouellette and Fraser (2009) found semantic effects in a similar experiment—but with consistent words (cf. Duff & Hulme, 2012, Experiment 2). In addition to these experimental studies, a number of correlational studies have shown that knowledge of

specific words' meanings relates to pronunciation accuracy (e.g., Kearns et al., 2016, in Grade 5 students; Goodwin, Gilbert, Cho, & Kearns, 2014, in middle school students; Nation & Cocksey, 2009, in 7-year-olds; Ricketts, Davies, Masterson, Stuart, & Duff, 2016, in 6-year-olds). The data overall appear to show that readers use semantic information about given words to support them when they read those specific words.

**General semantic information in word reading.** The complementary question is whether general semantic knowledge (for example, the size of a reader's lexicon) might also support reading of specific words—separate from word-specific knowledge. In a distributed processing account, this should be the case. Each encounter with a word leads to greater accuracy and greater fluency in subsequent encounters (Manis, 1985; Share, 1995). In a distributed processing account, however, each encounter improves the entire system, and these improvements occur both in the OP system as well as the OS system (Harm & Seidenberg, 2004; Rueckl, 2016). In an inconsistent orthography like English, the extent of a reader's general semantic knowledge might also be important. The OS system might be most helpful to readers with an extensive sound-meaning system (i.e., a good preschool vocabulary), while a poor vocabulary would result in less predictable OS mappings and therefore provide little benefit when reading aloud. Thus, we would expect a reader's general vocabulary to predict their word reading ability.

There are now multiple correlational studies to support this possibility (Jean & Geva, 2009, in Grade 5 and 6 students; Kearns et al., 2016, for irregular words; Mitchell & Brady, 2013, in Grade 4 students; Muter et al., 2004, in 6- and 7-year-olds; Ouellette & Beers, 2010, in Grade 1 and 6 children; Ricketts, Nation, & Bishop, 2007, in 8-, 9-, and 10-year-olds for irregular words; Ricketts et al., 2016, in 6-year-olds). The results show some ambiguity whether



general semantic effects should obtain for only irregular words, but these studies collectively suggest that readers with better semantic knowledge will have better word reading skills overall.

What no study has addressed is whether this general semantic knowledge is relevant separate from word-specific semantic knowledge. In the extant data, it is possible that a general semantic effect simply indexes the likelihood that a reader has knowledge of given word. For example, readers might pronounce *stirrup* correctly in part because of prior exposure to its spoken form, not because they know a lot of words in general. Disentangling these possibilities could provide indications how semantics might be integrated in word reading instruction, so the absence of this knowledge has instructional implications.

The present study separates item-specific and general semantic factors: We asked participants to read words aloud, ascertained whether they knew anything about the specific words' meanings, and measured the size of participants' vocabularies. We then analyzed the data using mixed-effects models that allowed us to examine whether the ability to read a word correctly related to knowledge of the specific word's meaning, the reader's overall semantic knowledge, or both.

### **Value of Semantic Knowledge in a Deep Orthography: Morphology**

The OS system is also relevant for word reading in that English frequently uses morphemes where letters reliably communicate information about their meanings, i.e., it is a (relatively) deep orthography (Frost & Katz, 1992). For example, *place* almost always involves the concept of location, and it has compound (*placemat*), inflected (*placed* or *places*), and derived (*replace* or *placement*) forms in which it retains its core meaning. The affixes *–ed*, *–s*, *re–*, and *–ment* also communicate meanings. The result is that the OS network could provide reliable semantic information at a sublexical level.

This would mean that a reader's morphological knowledge and skills would affect their word reading accuracy, and indeed they do: Morphological awareness—a measure of the ability to identify and manipulate morphological units in words—has a strong association with many reading tasks for children across the elementary grades, including word reading (e.g., Apel, Wilson-Fowler, Brimo, & Perrin, 2011; Deacon, 2012; Deacon & Kirby, 2004; Kirby et al., 2012; see Carlisle & Kearns, 2017, for a review). So, a reader's skill with morphological tasks could indicate the use of the OS pathway.

Another interpretation of the correlation between morphological awareness and word reading is that morphological awareness might index semantic rather than morphological knowledge (Spencer et al., 2017). However, several correlational studies have examined both morphological and semantic effects and found only the former (e.g., Kearns, 2015, in Grade 3 and 4 readers; Kearns et al., 2016, in Grade 5 students, and Goodwin et al., 2014, in Grade 7 and 8 students). These studies seem to indicate the importance of morphology over semantics—perhaps because morphology functions along OP and OS paths rather than just one. However, it is not clear whether a reader's general morphological skills, semantic knowledge, or both provide value when accounting for word-specific semantic knowledge. In this study, we included a measure of vocabulary size to index semantics and a morphological awareness measure requiring decomposition (e.g., changing *farm* to *farmer*). We selected these tasks to distinguish general semantic and morphological effects, but it is important to be clear that the latter task also has a semantic component. Whether the morphological measure shows an effect after accounting for vocabulary knowledge will provide some insight about the kind of meaning information readers might use to pronounce words.

However, morphology also involves OP connections (Gonnerman, Seidenberg & Anderson 2007; Katz & Frost, 1992; Nagy, Carlisle, & Goodwin, 2014). In other words, knowledge of a written word's morphemes relates to the word's meaning but also its pronunciation (e.g., *un-* is {not} and also /ʌn/). Because morphological skills integrate orthographic, semantic, and phonological knowledge at the sublexical level, readers could read morphologically complex words correctly using morphemes as OP units. Rueckl and Raveh (1999) observed that morphological information improved the functioning of the entire reading system. So, it may be that morphological effects relate to letter-sound connections. To control for this possibility in this study, we also included measures of general orthographic and phonological knowledge and skills.

### **The Current Study**

Overall, the data indicate a lack of knowledge about whether readers use knowledge of words' specific meanings or their vocabulary knowledge in general and whether general semantic effects are truly semantic or instead (or partly) morphological. No prior study has attempted to separate these effects, despite that different patterns would result in very different instructional implications: Should teachers teach the meanings of specific words they are learning to read (word-specific effect), focus on building children's overall vocabulary (general semantic effect), or focus on sublexical meaning units (general morphological effect).

The present study is designed to offer data toward understanding the role of semantics in word reading. We address these issues by examining third and fourth grade children's accuracy in reading polysyllabic words based on measurement of their word-specific semantic, orthographic, and phonological knowledge and measures of their general semantic, orthographic, phonological, and morphological skills. In other words, we attempt to isolate the nature of

semantic effects by measuring all parts of the distributed processing model of reading and determining which contribute to word reading accuracy. In doing this, we focus on polysyllabic words, owing to their increasing importance in the middle elementary grades (Kearns et al., 2016; Hiebert, Martin, & Menon, 2005); the complexity of their OP relationships (Arciuli, Monaghan, & Ševa, 2010; Cummins & Port, 1998; Goswami et al., 2011; Rastle & Coltheart, 2000); and the fact that morphologically complex words are almost always polysyllabic.

One unique feature of this study is that we use linear mixed-effects models to examine word-specific and general influences simultaneously (Baayen, Davidson, & Bates, 2008; Cho, Gilbert, & Goodwin, 2013). These models have become de rigueur in psychology, but their specification can be confusing. Figure 1 illustrates the model specification for the 48 words and 95 children in this study; additional details about the data structure are included in the online-only Appendices E and F. Our approach mirrors that of Ricketts et al. (2016). They also used linear mixed-effects models to understand the effect of item-specific semantic knowledge (termed “lexical semantics”) on six-year-olds’ word reading accuracy, and they controlled for word-specific phonological knowledge (termed “lexical phonology” by them) with an auditory lexical decision task. In a separate single-level regression model, they also measured the general effect of vocabulary. Our approach differs in that we test general and word-specific semantic effects in a single model and in that we ask whether general semantic effects are partly morphological.

The following models are presented to address the issues of interest: Model 1 concerns the association between a child’s semantic knowledge about a given word (word-specific knowledge) and accuracy in reading that word aloud (word-specific accuracy). Model 2 tests the general semantic effect. Model 3A combines the word-specific and general models to test both

types of effects. Model 3B included a morphological awareness variable to Model 3A to test whether semantic effects are conflated with morphological awareness.

## Method

### Participants

The participants were children with reading difficulty (RD;  $n = 46$ ) and their typically-achieving peers (TA,  $n = 49$ ). We over-sampled children with RD in middle elementary school because these children often receive beginning reading instruction of the kind described at the outset. Including a range of children allowed us to make claims across the range of achievement but provided greater precision in estimating the effects for children with RD. The students attended five different schools in the Northeastern United States, two serving mostly white English-speaking families of high socioeconomic status (SES) and three families with diversity in SES, ethnicity, and home language.

Group (TA/RD) assignment was based on readers' composite standard scores for the Sight Word Efficiency (SWE) and Phonological Decoding Efficiency (PDE) subtests of the *Test of Word Reading Efficiency, 2<sup>nd</sup> Edition* (TOWRE2; Torgesen, Wagner, & Rashotte, 2014). Children with RD had composite scores below the 25<sup>th</sup> percentile, and children with TA had composite scores above the 35<sup>th</sup> percentile. Table 1 presents the demographics for both groups.

### Measures

The outcome of interest was word reading for 48 polysyllabic words. We included both word-specific measures of semantic, orthographic, and phonological knowledge for the same 48 words as well as general measures of the same constructs. Table 2 provides descriptive statistics and bivariate correlations for these measures.

**Polysyllabic word reading.** Children read 48 polysyllabic words presented one at a

time, in random order, on a laptop computer screen using E-Prime 2.0 (Psychology Software Tools, 2014). Words appeared in 18-point Courier New font on a laptop with a 43 cm length-width diagonal. A fixation point (“+”) appeared on the screen for 1500 ms before the word appeared. The word stayed on the screen until the child responded or for 5000 ms. The examiner coded each response as correct or incorrect. E-Prime produced an audio file for each item. Later, two coders independently coded each response; Cohen’s  $\kappa$  was .93. Cronbach’s alpha was .90. The items appear in Appendix A.

Because all 95 children read all 48 words, the data were cross-classified, meaning that there was separate variability in responses due to the child’s ability and to the difficulty of each word. This data structure required us to use linear mixed effects models designed to permit appropriate specification of nested data. Figure 1 provides a visual representation of the specification of the models, and a schematic of the data structure appears in the online-only Appendix F, Figure 1. Further details about the models are included in the data analysis section below.

**Word-specific tests.** We administered tests of children’s orthographic, phonological, and semantic knowledge for each of the 48 words, that is, every child completed four tests related to the same 48 words, specifically, the word reading test and the three domain- and item-specific tests.

***Word-specific orthographic knowledge.*** For this, we used an orthographic choice task similar to that of Olson, Kliegl, Davidson, and Foltz (1985). Children were given four choices—a real word and three pseudohomophones; e.g., *renew* with *rinew\**, *renue\**, and *rinue\**—and asked to select the real word. Pseudohomophones were created by changing the first syllable (e.g., *symphony* to *simphony\**), the second or last syllable (e.g., *symfony\**), and a combination of

those two (*simfony*\*). Cronbach's alpha was .89. When we summed the items, the correlation with the Olson et al. task was .86.

***Word-specific phonological knowledge.*** Children's word-specific phonological knowledge was measured with a task that required them to listen to words pronounced without prosody with a 500 ms pause between each syllable. The child received a score of 1 for saying the correct word based on the stress-less pronunciation and a 0 otherwise. The internal consistency of this measure was .79. See Appendix B for details about the design of the measure.

***Word-specific semantic knowledge.*** Children's word-specific semantic knowledge was measured by presenting them with sentences—read aloud and shown on paper—that contained correctly used or misused target words. Students were asked to say whether each sentence made sense. The internal consistency was .82. When we summed the items, the correlation with performance on the Peabody Picture Vocabulary Test, 4<sup>th</sup> Edition (Dunn & Dunn, 2007) standard scores was .62. Appendix C includes further information about the test.

#### **General tests of child abilities.**

***Orthographic knowledge.*** Children's general orthographic knowledge was measured using an abbreviated version of the Olson et al. (1985) *Orthographic Choice Test*. Children were shown 65 pairs of words (a real word and a pseudohomophone; e.g., *tape* - *taip*) and asked to select the word that represented a real word. The original test had contained 80 items, but previous pilot data showed little variability in 15 items; these were removed to reduce length. Internal consistency for the 65-item measure was .94.

***Phonological knowledge.*** Children's general phonological knowledge was measured using four subtests from the Comprehensive Test of Phonological Processing (CTOPP; Wagner,

Torgesen, & Rashotte, 1999): Rapid Digit Naming (CTOPP-RDN), Rapid Letter Naming (CTOPP-RLN), Elision (CTOPP-E), and Nonword Repetition (CTOPP-NWR). They were used to capture the aspects of phonological processing described by Ramus and Szenkovits (2008).

**CTOPP-RDN.** Children were presented with a list of 36 numbers in 4 rows and asked to say each as fast as possible. The score was the number of seconds required to read this list. CTOPP-RDN has a test-retest reliability of .80 for children between ages 8 and 17.

**CTOPP-RLN.** Children were presented with a list of 36 letters, given like the CTOPP-RDN. The same scoring system was used. Test-retest reliability is .72 for children ages 8 to 17.

**CTOPP-E.** Children were asked to delete syllables or sounds from spoken words (e.g., “Say *doughnut* without saying *dough*.” and “Say *meat* without saying /t/.”). It has test-retest reliability of .79 for children ages 8 to 17, and internal consistency was .90 for this sample.

**CTOPP-NWR.** Children were asked to listen to a recorded presentation of increasingly complex nonwords (e.g., *joop*, *burloogajendaplo*) and repeat them exactly as they heard them. The test has test-retest reliability of .75 for children ages 8 to 17.

**Semantic knowledge.** General semantic knowledge was measured with the Peabody Picture Vocabulary Test, 4th Edition (PPVT-4; Dunn & Dunn, 2007). Children identified one of four pictures that represented a word spoken by the examiner. It has test-retest reliability of .93 for children between ages 7 and 10.

**Morphological awareness.** Morphological awareness was measured with Carlisle’s (2000) Test of Morphological Structure, Derivation subtest (TMS-D), using only the derivation subtest from that test (given that it correlated better than a decomposition subtest with reading in Carlisle’s study). The TMS-D measures children’s meta-linguistic ability to apply knowledge of derivational morphology by choosing a derived form of a word that fits the semantics and syntax



of a spoken sentence. Children completed 28 sentences by providing a derived word related to the given base word and the sentence context (e.g., “*Farm: My uncle is a \_\_\_\_*”). The examiner read the items aloud while the child read along on a printed version. This procedure assured children could rely on all possible aspects of lexical quality. The suffixes on the test were thought by Carlisle to be familiar to children in third through fifth grade.

Performance on the TMS-D is thought to capture two aspects of morphological awareness, morphological decoding and morphological analysis. It depends in part on knowledge of morphemes as commonly occurring phonological (and thus orthographic) units (decoding) and in part on knowledge of morphemes’ semantic (syntactic and conceptual) function (analysis), both of which support word pronunciation (Deacon, Tong, & Francis, 2017; Kuo & Anderson, 2006). In addition, performance on the TMS-D shows relations both with other measures of morphological awareness (e.g., Kruk & Bergman, 2013; Mitchell & Brady, 2014; Ramirez, Chen, Geva, & Luo, 2011) and with measures of vocabulary (e.g., Fracasso, Bangs, & Binder, 2016; Kruk & Bergman, 2013; Ramirez et al., 2011). Internal consistency for the TMS-D was .89 for this sample.

**Word characteristics.** Given our emphasis on semantic characteristics related to the child, we selected words based only on two word characteristics (cf. Goodwin et al., 2013; Goodwin, Gilbert, Cho, & Kearns, 2014; Kearns, 2015; Kearns et al., 2016), frequency and consistency (see Table 2 for descriptive statistics; see Appendix A, Table 1 for word-specific values). We included these only to assure that the word-specific effects could not be explained by word characteristics rather than child performance.

We included frequency because it indexes the likelihood that a word is in the child’s lexicon. We included grapheme-phoneme consistency (e.g., how frequently *ea* says /i/versus

anything else) because it has been shown to interact with frequency in studies of word recognition. Readers often show consistency effects only for low frequency words (Chateau & Jared, 2003; Jared, 2002; Jared & Seidenberg, 1990; Schmalz, Marinus, & Castles, 2013; Yap & Balota, 2009). For this reason, we included the frequency-consistency interaction.

We selected polysyllabic words with high/low frequency and consistency contrasts, including monomorphemic ( $n = 29$ ) and polymorphemic ( $n = 19$ ; all derivations) words to sample different types of polysyllabic words. To select the words, we began with unique words ( $N = 28,872$ ) in the CELEX English lemma database (Baayen, Piepenbrock, & van Rijn, 1993). We then limited the words to those in the *Educator's Word Frequency Guide* (Zeno, Ivens, Millard, & Duvvuri, 1995) with a frequency value in Grade 3 through Grade 8 lists ( $N = 4,822$ ). In other words, we eliminated words we expected children would never have seen or heard before. This was important because having words that would be completely unfamiliar to most readers would make this study a poor test of semantic effects. To select the high and low frequency words, we used the standard frequency index in the *Educator's Word Frequency Guide* to choose words that were higher frequency (at least 0.75 *SDs* above the mean) and lower frequency (at least 0.75 *SDs* below) among the available words.

For consistency, we calculated this variable—following other studies of polysyllabic words (Chateau & Jared, 2003; Jared & Seidenberg, 1990; Yap & Balota, 2009) rather than using regularity data (e.g., Rastle & Coltheart, 1999). Regularity data are typically based only on monosyllabic words. The type of consistency considered was between the *syllable body* and the *phonological rime*, similar to Glushko (1979). This was one type of consistency Yap and Balota (2009) showed it affected accuracy and response times for adult reading of polysyllabic words from the English Lexicon Project (ELP; Balota et al, 2007). There are many ways to calculate

consistency, and all require special considerations. Appendix D, Table 1 summarizes these considerations and shows the measure we chose, namely, *stressed-syllable body-rime feedforward type-based consistency*, following Jared and Seidenberg (1990) and Chateau and Jared (2003), mapping from the letters (the syllable body) to sound (the rime), following orthographic syllable division and based on types rather than tokens. The calculation for *pretty* is explained in Appendix D, Figure 1.

To determine whether this stressed-syllable body-rime metric was an appropriate metric for polysyllabic words, we ran exploratory analyses using lexical decision accuracy and response time data for all polysyllabic words in the ELP. We followed the seven-step process of Yap and Balota (2009) and compared this measure to others calculated by them. We opted for the stressed syllable consistency instead of one averaging all syllables because reduced vowels artificially reduce consistency. To separate the words by consistency, we used a 0.75 *SD* above/below cut-off for the words in the same *Educator's Word Frequency Guide* word set described above. In the analysis, the continuous frequencies and consistencies were used to take advantage of the additional variability and obtain more precise estimates of their influence.

## **Procedure**

**Test administration.** As described above, participants were screened using the TOWRE-2 SWE and PDE subtests. Only children meeting TA (TOWRE-2 > 35<sup>th</sup> percentile) or RD (TOWRE-2 < 25<sup>th</sup> percentile) criteria were included in the study. Test examiners worked with children in two separate sessions that lasted about 60 minutes. The sessions did not occur on the same day.

Administration order of the word-specific measures was important because children might profit on later tests from information obtained in earlier tests. Three principles governed

our placement of the word-specific tests: (1) tests would be ordered so earlier tests would not likely influence accuracy on later tests, (2) no word-specific test directly followed another word-specific test, and (3) the word-specific measures were spread across the two days to limit the possibility of carryover effects.

In the first session, we gave the orthographic choice task and the phonological knowledge task, and we gave the orthographic choice task first. The orthographic choice test procedure includes no support for pronouncing the words, but subvocalization of the items may have primed later phonological or word reading responses. It was decided that this had a lower likelihood of affecting word reading than the phonological task because it did not require pronunciation. The phonological knowledge task was administered after the orthographic knowledge task with one test between them. The phonological knowledge task had the potential to bias participants toward a particular response on the word reading test due to prior phonological exposure, but we judged that this was unlikely because the phonological stimuli were given as individual syllables without intonation rather than as whole words. In addition, the word reading test was given in a different session.

In the second session, the polysyllabic word reading test was administered before the semantic knowledge task. The word reading test had to precede the semantic task because the children were read the semantic items aloud and they could follow along with the printed sentences. Thus, children would have inflated word reading scores if they read the words after doing the semantic task.

Ordering the tests carefully and testing students on different days was a strategy used to prevent transfer effects. However, this does not eliminate the possibility of transfer across tasks, but it limited it as much as possible given that the purpose of the study necessitated the use of

multiple measures of knowledge about the same words.

**Data validation.** All measures were scored and double-scored by the examiners, trained by the authors. After test practice, test administration fidelity, and scoring accuracy were evaluated and had to be at least 90% before testing children. After testing, scores were double-checked, entered, double-entered, and checked for discrepancies using REDCap (Harris et al., 2009).

### **Data Analysis**

The main analytical tool was a linear mixed-effects model, also called an explanatory item response model (De Boeck & Wilson, 2004). These models include random person (child) and item (word) parameters (Janssen, Schepers, & Peres, 2004; van den Noortgate, De Boeck, & Meulders, 2003). Analyses were conducted using the lmer function (Bates et al., 2015) from the lme4 library in R. The models explain variation in the probability that a response would be correct given variability in person abilities (child skills) and item difficulties (word characteristics) with item-specific fixed effects (child-specific scores on word-specific measures) while accounting for classroom and school random effects. For readers familiar with multilevel models (e.g., Raudenbush & Bryk, 2002), this is equivalent to a 4-level model with crossed random effects (person and item) and a binomial outcome (polysyllabic word reading accuracy). More specifically, responses at level 1 are cross-classified with persons and items at level 2, nested within classrooms at level 3 and schools at level 4. Figure 1 shows how the models account for child and word characteristics using a specific example. Details concerning the specification of the models are provided in the online-only Appendix E.

These models have two benefits over ordinary least-squares regression (Jaeger, 2008). First, these permit us to simultaneously estimate child and word effects without the need to

present separate tests for persons and items (i.e., *F1* and *F2*) as was long common practice in psychology and linguistics (Clark, 1973; Raaijmakers, Schrijnemakers, & Gremmen, 1999). Second, because data do not need to be collapsed across persons or items, this approach provides greater power for statistical tests without assuming that the data reflect a random sample of children and words (an assumption possibly increasing Type I error), as simulations by Stevens and Brysbaert (n.d.) have shown. Although there are not power analyses available for linear mixed-effects models, Stevens and Brysbaert found that these models could detect effects that were less than 20% of the standard deviation of the random person effect with 80 participants and only 20 items with about 70% power. Given that we had more participants and items, we judged that the study included a sufficient number of participants to have adequate power to detect effects even if they were relatively modest. For the analyses, all continuous variables have been *z*-scored to make the coefficients interpretable in standard deviation units.

## Results

The descriptive data for the word-specific measures provide support for the idea that semantic knowledge and polysyllabic word reading skill are linked. Table 3 provides crosstabulations of correct and incorrect responses ( $N = 4,393$ ) for the polysyllabic word reading outcome and the word-specific semantic, orthographic, and phonological tasks. The large  $N$  reflects that we recorded every child's response for every word, allowing us to examine how responses for each orthographic, phonological, and semantic item relates to reading accuracy for the specific word, as described in the Data Analysis section.

Table 3 also shows the strong alignment between children's semantic knowledge of polysyllabic words and their ability to read them correctly. The table shows that there was perfect alignment between performance on the two semantic items 49% of the time, 62% if

getting one item in the pair correct counted as incorrect. This is well above the chance probability. Based on the analysis of the responses in Table 3, we created a dichotomous semantic knowledge variable by counting both incorrect and one incorrect responses as incorrect, bringing it in line with the dichotomous coding for orthographic and phonological knowledge. For the other variables, word reading accuracy and phonological knowledge matched in 67% of cases and orthographic knowledge in 69%. There was enough variability in the word reading outcome and item-specific covariates to test semantic effects.

### **Unconditional and Control Models**

Before estimating the models of interest, we established the necessary random effects structure using an unconditional—or null—model and identified the need to include random effects for child word, classroom, and school. The details of model specification are given in the online-only Appendix E. We calculated the range of *plausible values* for children and words, that is, the range in which 95% of probabilities of a correct response are expected to occur, to decide if there was sufficient variability in responses to run subsequent models. The child range was .33 to .93, indicating where 95% of responses are expected to fall—for a word of average difficulty. The word range was .20 to .96 for a child of average ability. Child and word performance varied considerably, so modeling the effects of our independent variables was appropriate.

### **Word-Specific Semantic Knowledge**

Model 1 was a test of the proposition that word-specific semantic knowledge, in terms of simple definitional knowledge as assessed by our semantic task, affected the polysyllabic word reading accuracy for the children in this study. The tested model also included word-specific experimental orthographic and phonological knowledge measures, to assure that we measured all

dimensions of the distributed processing model and to avoid the possibility that semantic effects were spurious. When the word-specific model was tested against the control model, the control model fit worse than that with the word-specific variables,  $\Delta\chi^2_{[3]} = 45.891, p < .0001$ .

The results of the correctly specified model are given in the first section of Table 4. Table 4 also includes information about the standard errors and statistical significance of the models, so these are not reported here. Coefficients are given in log-odds to permit the non-linear outcome—probability of a correct response—to be measured as a linear function. It is easier to interpret probabilities but they are not linear. When log-odds are changed to probabilities, their effects appear smaller when probabilities are near 1 or 0 than if the probability is closer to .50. Table 4 presents the coefficients in log-odds with accompanying probabilities, and Table 5 provides a guide for understanding the relationship between them.

For Model 1, the intercept indicated that probability of reading a word correctly was .08 for a third grader with RD and incorrect responses on all three word-specific items. The probability was .11 if a child answered the semantic knowledge items correctly, .12 when the correct orthographic choice was made, and .12 probability when the correct phonological response was given. If the child got only the semantic item wrong, the probability of a correct response was .17. If all three items were correct, the probability was .24. All three effects were similar in magnitude and statistically significant, indicating that word-specific semantic effects were independent of orthographic and phonological ones.

### **General Semantic Knowledge**

Model 2 was used to examine whether there was a general semantic knowledge effect, based on a measure of vocabulary breadth. We included general orthographic and phonological knowledge measures to address the possibility that a general semantic effect would not be



present when we controlled for these other two aspects of lexical quality. To specify Model 2, we added the fixed effects for the  $z$ -scored general measures and then determined whether random word slopes were necessary for correct model specification. The model containing only the fixed effects for the general child variables fit worse than one containing random word slopes for PPVT-4 and orthographic choice,  $\Delta\chi^2_{[2]} = 39.702, p < .001$ .

For the correctly specified model, three effects of general skills were observed, reflecting that all distributed processing skills contributed to word reading accuracy. As with Model 1, all probabilities are based on the response of a third grader with RD with mean scores on all five general measures. The intercept probability was .48. A 1 *SD* increase in the child's score on vocabulary as measured by PPVT, orthographic knowledge as measured by the Olson et al. (1985) orthographic choice task, or phonological awareness as measured by CTOPP Elision increase the probability of a correct response to .54, .55, or .54 respectively. A 1 *SD* decrease from the mean score would decrease it to .42, .41, or .42, respectively, when scores on all other variables were at their means. These differences were large enough to support the idea that general semantic knowledge has a practically meaningful relationship with word reading skill.

### **Word-Specific and General Semantic Effects**

Model 3A is the key test of the question whether semantic effects reflect children's word-specific knowledge or their general semantic knowledge. For this test, we combined the word-specific knowledge variables from Model 1 and the general variables from Model 2. The model without a random word slope for PPVT still fit worse than the model with it,  $\nu_{030} = 0.261, \Delta\chi^2_{[1]} = 6.615, p = .01$ , as did the model without the random word slope for general orthographic choice,  $\nu_{040} = 0.334, \Delta\chi^2_{[1]} = 16.264, p < .0001$ . The intercept probability was .29 for a third grader with RD with mean scores on all general measures and incorrect scores

on all three word-specific measures. If the semantic item was correct, probability of a correct response was .37. Figure 2, top left, shows this difference. The effects for correct orthographic and phonological responses are also shown for comparison.

All three general child skills also continued to relate to polysyllabic word reading accuracy. Vocabulary, orthographic knowledge, and phonological awareness scores 1 *SD* above the mean had predicted probabilities of correct responses of .25, .29, and .33, holding all other variables to the intercept (see Figure 2, bottom).

### **Word-Specific and General Semantic Effects including Morphological Awareness**

Model 3B contained the same effects as the prior model, plus the morphological knowledge effect. The model without morphological knowledge fit worse than Model 3A,  $\Delta\chi^2_{[1]} = 9.059$ ,  $p = .003$ . A child who got the semantic item correct had a significantly higher probability of a correct response than a child who did not (.53 vs. .43), holding all else constant (see Figure 3). Among the general effects, only phonological awareness and morphological awareness as measured by the TMS-D had significant effects. Neither the vocabulary nor the orthographic knowledge effect was reliable. The morphological awareness effect, based on TMS-D scores 1 *SD* below the mean, at the mean, and 1 *SD* above the mean, resulted in predicted probabilities of .34, .43, and .53, all else held constant. Figure 3, bottom, shows this.

### **Discussion**

To review the focus of this study, distributed processing models of English word reading suggest that semantic effects on word pronunciation are possible, on the basis that orthographic-semantic (OS) pathways can support pronunciation when orthographic-phonological (OP) information is ambiguous, as it is often with polysyllabic words. However, tests of distributed processing models (e.g., Harm & Seidenberg, 2004) have not yet shown whether readers use

semantic information to facilitate word pronunciation. Theory suggests it is possible, and some extant data support this idea. However, semantic information might operate on specific words, in general, or in the form of morphology. This study addresses all three possibilities.

To summarize the results, we tested word-specific semantic knowledge controlling for word-specific phonological and orthographic knowledge, finding effects of similar magnitude for all three. We tested general semantic effects controlling for general orthographic and phonological knowledge, and found effects of similar magnitude for all three. When we combined word-specific and general effects, the word-specific and general semantic effects remained reliable. However, when we included a measure of morphological awareness in the analysis, the word-specific semantic effect was present but the general vocabulary effect as measured by the PPVT4 was not (nor was there an orthographic effect).

### **Semantic Knowledge in a Distributed Processing Account of Word Reading**

The present study provides some support for the rapidly-advancing idea that semantic knowledge has an important role in word reading (Taylor et al., 2015). We consider these findings robust on the basis that we controlled for other aspects of a distributed processing system—orthographic and phonological knowledge—that account for a great deal of variability in word reading in the extant literature.

**Item-specific semantic knowledge.** The presence of an item-specific semantic knowledge effect is particularly convincing because we included measures of phonological and orthographic knowledge that could easily have consumed most of the variance at the word level. It is also perfectly consonant with the word-specific finding of Ricketts et al. (2016), even though the semantic tasks were different (deciding if sentences made sense in our case; giving definitions in theirs). It also adds strength to the Ricketts et al. results because of the controls for

orthographic and phonological knowledge at the general and item-specific levels. The nature of the item-specific task used here required knowledge of words' definitions, so the effect affects readers' conceptual knowledge of words affect their word reading performance when they encounter these words in texts.

There may be a second form of item-specific semantic knowledge, knowledge simply that a word is in the lexicon, what Ricketts et al. (2016) called "lexical phonology." The evidence comes from the item-specific phonological effect. The aggregated scores on the phonological knowledge had a relatively high correlation with semantic knowledge ( $r = .57$ ), higher than that for the orthographic knowledge composite ( $r = .38$ ). The phonological task required readers to make minor adjustments to words and produce an accurate pronunciation. It may be that readers adjust these pronunciations to find entries in the phonological lexicon, similar to the idea put forth by Kearns, Rogers, Al Ghanem, and Koriakin (2016) that readers have a semantic and phonological ability to adjust recoding (SPAAR) when they recode words incorrectly—based on earlier similar results from Elbro et al. (2012) and Tunmer and Chapman (2012). The present study suggests that SPAAR operates as Kearns et al. suggested at the item level. In sum, the item-specific semantic effects involve both conceptual knowledge of words and knowledge that the words are lexical entries.

**General semantic knowledge.** We did not observe a semantic effect based on performance on a measure of vocabulary breadth (PPVT4) when we considered morphological awareness as measured by the TMS-D. However, the TMS-D is not a pure measure of morphological awareness. Studies by Carlisle and Fleming (2003), Fracasso et al. (2016), Kruk and Bergman (2013), and Ramirez et al. (2011) found strong correlations between scores on TMS-D performance and on vocabulary and reading comprehension tests. Perhaps more

important are data from Spencer et al. (2015) showing that multiple measures of morphological awareness and of vocabulary knowledge load onto a single morphological/vocabulary knowledge factor that included the PPVT4, the Stanford-Binet Vocabulary subtest, the TMS-D, and other measures of morphological awareness (e.g., Nagy, Berninger, Abbott, Vaughn, & Vermeulen, 2003). In other words, the effect of TMS-D we observed is—at least in part—evidence that general semantic knowledge is relevant for word reading. The strong relationship between the TMS-D and general vocabulary results from the fact that the task requires children to locate lexical entries that contain the base word, so those who have better vocabulary skills are more likely to know the base words and their derivation that appear on the TMS-D. This general semantic effect aligns with a growing body of literature suggesting that semantic knowledge affects word recognition, even in isolation, as measured here (see Taylor et al., 2015, for an overview). In distributed processing terms, the size of the reader's semantic network affects individual word reading.

As important, the relationship between word reading and morphological awareness skill is not only semantic: Morphemes certainly have OS links, but they also connect orthographic and phonological units. Morphological units provide redundancy across OS and OP units that simultaneously support the production of the correct spoken word. On the OS side, exposure to morphemes means that readers develop sensitivity to the relationship between concepts and letters that would result in a stable OS system that supports word recognition when the OP connections for morphemes are less useful (e.g., *nature* vs. *natural*). On the OP side, morpheme exposure can stabilize the letter-sound system by associating longer strings of letters with larger groups of phonemes. This will support OP processing even if the morphemes have no semantic value (e.g., *cran-* in *cranberry*). This idea aligns with the lexical quality hypothesis, where a

reader's representations of words' orthographic, phonological, and semantic units differ in representation quality (Perfetti, 2007). Morphemes may function similarly. Readers may have different levels of OP and OS knowledge for morphemes, meaning that their value for readers depends on the degree to which they provide both types of information. Morphemes for which readers have strong OP and OS connections are likely to support polymorphemic word recognition best. For example, children in the late elementary grades likely have strong OP and OS representations of *-less*, but they may know the pronunciation of *-ity* better than its semantic and syntactic function. Accordingly, we would say that words with *-less* are more tractable than those with *-ity*, all else being equal.

In addition, the morphological awareness finding continues to suggest that the size of a reader's vocabulary also matters—even when accounting for word-specific knowledge. Part of what makes it possible for children to succeed on the TMS-D task is their knowledge of the specific base words and derived words on the test. For example, children who say *equalness* for *equality* probably lack a lexical entry for the latter. What is important about this result in the present study is that this general effect is not standing in for an item-specific effect—because we explicitly measured item-specific knowledge.

There are two complementary explanations for the presence of a semantic-morphological effect. In terms of semantics, one possible explanation for a general effect is that readers with larger vocabularies persist in the search for a lexical entry when it cannot be immediately retrieved, a possibility suggested by Kearns (2015) and Ricketts et al. (2007). In terms of morphology, the effect may indicate that—when the OP system is sensitive to morphology—readers can use OP morpheme knowledge to produce plausible (and therefore often correct) pronunciations of unknown words. By contrast, when the OP system does not account for

morphology, readers may need to apply knowledge of smaller units to read unknown words (e.g., grapheme-phoneme correspondences like in *-ment* with four letter-sounds but one morpheme). Relying on small units alone will require more effort than using small units and morphemes because this requires keeping more information in working memory. In addition, small units are often inconsistent and likely to require semantic and phonological ability to adjust recoding to produce a correct response (Kearns, Rogers, Koriakin, & Al Ghanem, 2016; Treiman, Mullenix, Bijeljac-Babic, & Richmond-Welty, 1995). Therefore, readers without morphological awareness may have a lower likelihood of producing the correct answer for words without lexical entries. This may also precipitate lower task persistence, as with the semantic effect. Overall, the data indicate that children's ability to read polysyllabic words depends on their general semantic and morphological knowledge and not simply their word-specific semantic knowledge.

### **Orthographic Knowledge (Non)Effects**

The purpose of the present study was not to examine whether orthographic knowledge is general or word-specific. However, researchers have debated the interpretation of data showing readers have a general orthographic processing skill that is separable from their knowledge of individual words' spellings (cf. Burt, 2006; Cunningham, 2006; Cunningham, Perry, & Stanovich, 2001; Cunningham & Stanovich, 1990; Deacon, 2012; Vellutino, Scanlon, & Tanzman, 1994; Wagner & Barker, 1994). It is worth noting that the present study offers some support for the idea that orthographic knowledge is word-specific because general orthographic knowledge effects were not present in our final model. However, we did find a general orthographic effect when we did not include the morphological awareness measure.

It is possible that orthographic knowledge effects—in general—are related to morphological awareness. This certainly aligns with the idea that morphological effects are

semantic and orthographic (Gonnerman et al., 2007). Perhaps researchers could examine more closely the idea that orthographic processing skill and morphological awareness are strongly connected abilities that are different dimensions of a meaningful general orthographic skill. The present study may serve as a useful starting point for future studies of orthographic processing. This paper offers one possibility for addressing this controversy, to test in the same statistical model both word-specific and general orthographic knowledge and their associations with word reading.

### **Educational Implications**

We began this paper with the observation that most beginning reading instruction in English focuses on developing orthographic and phonological knowledge. However, a distributed processing account of word reading makes it possible that OS connections could affect word reading, raising questions whether programs to teach beginning word reading should include emphasis on semantic aspects of words. Although these cross-sectional data are not causal, they provide the opportunity to consider the potential value of semantic instruction.

**Teaching word reading via semantics and morphology.** Before addressing the instructional question, it is important to establish whether the semantic effects are practically meaningful. In the final model, we saw evidence for practical significance. At the word-specific level, the probability of a correct word reading response increased from .43 to .53 if the word-specific semantic knowledge question was answered correctly. In terms of general semantic knowledge and morphological awareness, having a score 1 *SD* above the mean on the morphological awareness measure also increased the score from .43 to .53. Both item-specific and general semantic knowledge could change the direction of a decision when the probability of a correct response is between .40 and .60. Moreover, the morphological awareness score was



larger than the phonological awareness effect—and phonological awareness instruction is widely agreed to be important (Adams, 1990). We regard the effects as practically meaningful.

In terms of word-specific semantic knowledge, the results suggest that children should practice reading skills with words likely in their phonological lexicon, particularly given data suggesting that readers adjust recodings using semantic and phonological information (Kearns et al., 2016; Elbro et al., 2012; Tunmer & Chapman, 2012). To wit, Dyson, Best, Solity, and Hulme (2017) and Savage, Georgiou, Parrila, & Maiorano (2018) found that teaching children to adjust recoding improved their overall reading skills. In other words, reading unknown words involves a search of the phonological lexicon for similar entries. These data support the importance of considering familiarity in choosing words for word reading instruction, turning away from arcane words like *pipkin*, a word at least one curriculum includes in their instructional materials as a real word. In light of this, it is noteworthy that early-reading curricula have gradually moved toward including less familiar words over the past six decades (Fitzgerald et al., 2016). The present data indicate the shift toward less familiar words may have the unintended consequence of making it harder for readers to build skill in using semantic information in word recognition.

At the same time, the taught words need not be very familiar. Distributed processing models make clear that even very limited knowledge about a word's meaning can contribute to correct pronunciation, so even a small amount of meaning information is helpful. Teacher might help students by pronouncing and briefly defining words they will practice at the beginning of a lesson, perhaps exploiting the fast mapping mechanism that results in vocabulary learning following very brief, simple exposure (Biemiller & Boote, 2006). Moreover, this is not an argument against reading nonwords—an activity in some reading programs. Nonword reading

practice refines readers' ability to use OP connections to read words, but this is insufficient because it does not support the ability to adjust recoding to locate a real word.

Turning to the general semantic knowledge result, the implications are probably very simple: First, children need to develop broad and deep vocabulary knowledge to increase the likelihood they will persist in the recoding process and to simultaneously facilitate reading comprehension (Catts, Herrera, Nielsen, & Bridges, 2015; Elleman, Lindo, Morphy, & Compton, 2009; Kendeou, Savage, & van den Broek, 2009; Ricketts et al., 2007). Time spent learning word meanings and engaging in meaning-focused activities could have an indirect effect on word reading. Second, the morphological aspect of the finding suggests that children might benefit from practice reading words' morphological units—building OP knowledge of them. It is also likely useful for students to learn about morphemes' semantic and syntactic functions—building OS knowledge. However, we would also argue that this is not the same as telling students' morphemes' definitions. Knowing that *-tion* means “the act of” is probably less valuable than discussing the meanings of sentences where the *-tion* word has an important function. For example, students might explain a sentence like “The teacher’s actions prevented the student from falling off the swing,” in their own words, perhaps saying something like, “This means that the teacher acted, did something that helped the student and stopped her from falling.” In this example, the student shows a developing representation of *-tion* by saying (a) “the teacher acted” and using the base word *act* as a verb with the past tense inflection and (b) “did something” indicates that the student knows the “act of” function of *-tion*. The student would not need to know the abstruse meaning of *-tion* in order to benefit from this kind of practice.

However, it is unclear (but important to establish) what kind of morphological practice activities might best support reading because the alternatives differ considerably in what aspects

of morphology they exploit. Currently available programs range considerably in the degree to which they emphasize morphemes as OP units (e.g., Lindamood & Lindamood, 1998; Lovett, Lacerenza, & Borden, 2000) or OS units (Bowers & Kirby, 2010; Crosson, McKeown, Moore, & Ye, 2018; Henry, 1989). In terms of resolving the question which activities in these programs are best, these results are helpful in two ways. First, children's word recognition skills will benefit from instruction on both OP and OS functions of morphemes because the units seem to have value in both ways. Second, children may benefit most from activities that contain extensive practice. In a distributed processing model, the stability of OP and OS knowledge improves most through extensive exposure (i.e., via statistical learning). So, it may be more profitable to do extensive practice rather than more cognitively demanding and time-intensive activities like deducing the pronunciations or meanings of affixes. That said, this study does not include an evaluation of the value of practice, so further discussion on this point is beyond the scope of the paper.

### **Limitations**

**Measurement.** One failing of all correlational studies is the possibility of endogeneity, i.e., omitted-variable bias. We used experimental measures and cannot know whether the measures we designed were ideal for measuring the targeted constructs. The phonological task involved syllable blending, for example. This is just one possible index of phonological knowledge. Another might have been the auditory lexical decision task used by Nation and Cocksey (2009) and Ricketts et al. (2016). Put simply, experimental tasks always raise the specter of incomplete and/or imprecise measurement.

In addition, children were given several tests with the same words, so a priming effect is a concern. For example, on the word-specific orthographic choice task, children likely

subvocalized the pronunciation of the items, priming a correct response for the word reading or phonological test given later and inflating scores on those tests. Differences on the word reading task could inflate scores but would not change the pattern of effects unless there was a ceiling effect (there was not). It could change interpretation if orthographic and phonological scores interacted; a post-hoc test of this possibility did not produce a significant effect ( $p = .63$ ).

We also selected specific words by frequency and consistency and did not contrast performance on the polysyllabic words with performance on monosyllabic words—words that have multiple potentially important differences (Kearns et al., 2016). Our concern is reduced by the fact that Ricketts et al. (2016) list included both word types and had results consistent with these.

In addition, we over-sampled students with RD. We wanted to be precise about the effects on this population given that these children often receive supplemental word reading instruction in middle elementary school. This may over- or under-emphasize the importance of semantics because there are more of these children than there would be in a normally-distributed sample. In a separate analysis, we examined whether the semantic effects differed as a function of TA/RD status, and they did not. Nonetheless, our sample selection procedure is a limitation.

### **Conclusion**

Distributed processing models of English word recognition suggest that OS connections may help students pronounce words, largely due to OP inconsistencies in English. Unresolved are whether semantic information might operate on specific words, be useful in general, or support readers primarily in the form of morphology. We tested these possibilities and found that the ability to read words was associated with readers' knowledge of the specific words' meanings, children's general morphological awareness, and (to a limited extent) their general

semantic knowledge. The data support the idea that word reading is a distributed process with OS knowledge affecting word reading in conjunction with (not instead of) OP knowledge. In terms of education, we think that developers of beginning reading curricula should consider (a) avoiding using very rare words when students are practicing decoding so they can use their distributed knowledge to support word reading, (b) focusing on words' morphological features—both as orthographic and semantic units—relatively early in word recognition instruction, and (c) should provide extensive practice that will build strong connections between children's representations of words spellings and the associated pronunciations and meanings.

## References

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: Massachusetts Institute of Technology Press.
- Anthony, J. L., Lonigan, C. J., Burgess, S. R., Driscoll, K., Phillips, B. M., & Cantor, B. G. (2002). Structure of preschool phonological sensitivity: Overlapping sensitivity to rhyme, words, syllables, and phonemes. *Journal of Experimental Child Psychology*, 82, 65-92. doi:10.1006/jecp.2002.2677
- Apel, K., Wilson-Fowler, E. B., Brimo, D., & Perrin, N. A. (2011). Metalinguistic contributions to reading and spelling in second and third grade students. *Reading and Writing*, 25, 1283-1305. doi:10.1007/s11145-011-9317-8
- Arciuli, J., Monaghan, P., & Ševa, N. (2010). Learning to assign lexical stress during reading aloud: corpus, behavioral, and computational investigations. *Journal of Memory & Language*, 63, 180-196. doi:10.1016/j.jml.2010.03.005
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412. doi: 10.1016/j.jml.2007.12.005
- Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX Lexical Database. Philadelphia, PA: Linguistic Data Consortium.
- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39, 445-459. doi:10.3758/BF03193014
- Bates, D. M., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Dai, B. (2015). Linear mixed-effects models using Eigen and S4 (Version 1.1-7).

- Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, 98, 44-62. doi:10.1037/0022-0663.98.1.44
- Bowers, P. N., & Kirby, J. R. (2010). Effects of morphological instruction on vocabulary acquisition. *Reading and Writing*, 23, 515-537. doi:10.1007/s11145-009-9172-z
- Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, 80, 144-179. doi:10.3102/0034654309359353
- Burt, J. S. (2006). What is orthographic processing skill and how does it relate to word identification in reading? *Journal of Research in Reading*, 29, 400-417. doi:10.1111/j.1467-9817.2006.00315.x
- Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing*, 12, 169-190.
- Carlisle, J. F., & Fleming, J. (2003). Lexical processing of morphologically complex words in the elementary years. *Scientific Studies of Reading*, 7, 239-253. doi:10.1207/S1532799XSSR0703\_3
- Catts, H. W., Herrera, S., Nielsen, D. C., & Bridges, M. S. (2015). Early prediction of reading comprehension within the simple view framework. *Reading and Writing*, 28, 1407-1425. doi:10.1007/s11145-015-9576-x
- Chateau, D., & Jared, D. (2003). Spelling-sound consistency effects in disyllabic word naming. *Journal of Memory and Language*, 48, 255-280. doi:10.1016/S0749-596X(02)00521-1
- Cho, S. J., Gilbert, J. K., & Goodwin, A. P. (2013). Explanatory multidimensional multilevel

- random item response model: An application to simultaneous investigation of word and person contributions to multidimensional lexical representations. *Psychometrika*, 78, 830-855.
- Cho, S.-J., Partchev, I., & De Boeck, P. (2012). Parameter estimation of multiple item response profile model. *British Journal of Mathematical & Statistical Psychology*, 65, 438-466.  
doi:10.1111/j.2044-8317.2011.02036.x
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning & Verbal Behavior*, 12, 335-359.  
doi:10.1016/S0022-5371(73)80014-3
- Clements, G. N., & Keyser, S. J. (1983). *CV phonology: A generative theory of the syllable* (Vol. 9). Cambridge, MA: The MIT Press.
- Crosson, A. C., McKeown, M. G., Moore, D. W., & Ye, F. (2018). Extending the bounds of morphology instruction: teaching Latin roots facilitates academic word learning for English Learner adolescents. *Reading and Writing*. doi:10.1007/s11145-018-9885-y
- Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics*, 26, 145-171.
- Cunningham, A. E. (2006). Accounting for children's orthographic learning while reading text: do children self-teach? *Journal of Experimental Child Psychology*, 95, 56-77.  
doi:10.1016/j.jecp.2006.03.008
- Cunningham, A. E., Perry, K. E., & Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Reading and Writing*, 14, 549-568.
- Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of*



- Educational Psychology*, 82, 733-740. doi:10.1037/0022-0663.82.4.733
- De Boeck, P., & Wilson, M. (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. New York, NY: Springer-Verlag.
- Deacon, S. H. (2012). Sounds, letters and meanings: the independent influences of phonological, morphological and orthographic skills on early word reading accuracy. *Journal of Research in Reading*, 35, 456-475. doi:10.1111/j.1467-9817.2010.01496.x
- Deacon, S. H., & Kirby, J. R. (2004). Morphological awareness: Just “more phonological”? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics*, 25, 223-238.
- Deacon, S. H., Tong, X., & Francis, K. (2017). The relationship of morphological analysis and morphological decoding to reading comprehension. *Journal of Research in Reading*, 40, 1–16. doi: 10.1111/1467-9817.12056
- Duff, F. J., & Hulme, C. (2012). The role of children's phonological and semantic knowledge in learning to read words. *Scientific Studies of Reading*, 16, 504-525.
- Dunn, L., & Dunn, D. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Pearson, Minneapolis, MN.
- Dyson, H., Best, W., Solity, J., & Hulme, C. (2017). Training mispronunciation correction and word meanings improves children’s ability to learn to read words. *Scientific Studies of Reading*, 1-16.
- Elbro, C., de Jong, P. F., Houter, D., & Nielsen, A.-M. (2012). From Spelling Pronunciation to Lexical Access: A Second Step in Word Decoding? *Scientific Studies of Reading*, 16, 341-359.
- Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The Impact of Vocabulary

- Instruction on Passage-Level Comprehension of School-Age Children: A Meta-Analysis. *Journal of Research on Educational Effectiveness*, 2, 1-44.
- Fracasso, L. E., Bangs, K., & Binder, K. S. (2016). The contributions of phonological and morphological awareness to literacy skills in the adult basic education population. *Journal of Learning Disabilities*, 49, 140-151.  
doi:10.1177/0022219414538513
- Fitzgerald, J., Elmore, J., Relyea, J. E., Hiebert, E. H., & Stenner, A. J. (2016). Has first-grade core reading program text complexity changed across six decades?. *Reading Research Quarterly*, 51, 7-28. doi:10.1002/rrq.115
- Frost, R., & Katz, L. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Advances in psychology: Orthography, phonology, morphology, and meaning* (Vol. 94; pp. 67–84). Amsterdam, Netherlands: North-Holland. doi:10.1016/S0166-4115(08)62789-2
- Gibson, E. J. (1965). Learning to Read. *Science*, 148, 1066-1072. doi:10.2307/1716844
- Gilbert, J. K., Compton, D. L., & Kearns, D. M. (2011). Word and person effects on decoding accuracy: A new look at an old question. *Journal of Educational Psychology*, 103, 489–507. doi:10.1037/a0023001
- Gillingham, A., & Stillman, B. W. (2014). *The Gillingham manual: Remedial training for students with specific disability in reading, spelling, and penmanship* (8<sup>th</sup> ed.). Cambridge, MA: Educators Publishing Service.
- Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 674-691. doi:10.1037/0096-1523.5.4.674

- Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General*, 136, 323-345. doi:10.1037/0096-3445.136.2.323
- Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, 17, 257-285. doi:10.1080/10888438.2012.689791
- Goodwin, A. P., Gilbert, J. K., & Cho, S. (2013). Morphological contributions to adolescent word reading: An item response approach. *Reading Research Quarterly*, 48, 39-60.
- Goswami, U., Porpodas, C., & Wheelwright, S. (1997). Children's orthographic representations in English and Greek. *European Journal of Psychology of Education*, 12, 273-292. doi:10.1007/BF03172876
- Goswami, U., Wang, H. L. S., Cruz, A., Fosker, T., Mead, N., & Huss, M. (2011). Language-universal sensory deficits in developmental dyslexia: English, Spanish, and Chinese. *Journal of Cognitive Neuroscience*, 23, 325-337. doi:10.1162/jocn.2010.21453
- Gough, P. B., Juel, C., & Griffith, P. L. (1992). Reading, spelling, and the orthographic cipher. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 35-48). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009) Research electronic data capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42, 377-81.
- Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia:

- insights from connectionist models. *Psychological Review*, 106, 491.
- Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review*, 111, 662.
- Hiebert, E. H., Martin, L. A., & Menon, S. (2005). Are there alternatives in reading textbooks? an examination of three beginning reading programs. *Reading & Writing Quarterly*, 21, 7–32. doi:10.1080/10573560590523621
- Henry, M. K. (1989). Children's word structure knowledge: Implications for decoding and spelling instruction. *Reading and Writing*, 1, 135–152. doi:10.1007/BF00377467
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434–446. doi:10.1016/j.jml.2007.11.007
- Janssen, R., Schepers, J., & Peres, D. (2004). Models with item and item group predictors. In P. De Boeck & M. Wilson (Eds.), *Explanatory item response models: A generalized linear and nonlinear approach*. New York, NY: Springer-Verlag.
- Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language*, 46, 723–750. doi:10.1006/jmla.2001.2827
- Jared, D., & Seidenberg, M. S. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 92–105. doi:10.1037/0096-1523.16.1.92
- Jean, M., & Geva, E. (2009). The development of vocabulary in English as a second language children and its role in predicting word recognition ability. *Applied Psycholinguistics*, 30, 153–185. doi:10.1017/S0142716408090073

- Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. *Advances in psychology*, 94, 67-84.
- Kearns, D. M., Rogers, H. J., Al Ghanem, R., & Koriakin, T. (2016). Semantic and phonological ability to adjust recoding: A unique correlate of word reading skill? *Scientific Studies of Reading*. doi:10.1080/10888438.2016.1217865
- Kearns, D. M. (2015). How elementary-age children read polysyllabic polymorphemic words. *Journal of Educational Psychology*, 107, 364–390. doi:10.1037/a0037518
- Kendeou, P., Savage, R., & van den Broek, P. (2009). Revisiting the simple view of reading. *British Journal of Educational Psychology*, 79, 353-370.  
doi:10.1348/978185408X369020
- Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing*, 25, 389-410. doi:10.1007/s11145-010-9276-5
- Kruk, R. S., & Bergman, K. (2013). The reciprocal relations between morphological processes and reading. *Journal of Experimental Child Psychology*, 114, 10-34.  
doi:10.1016/j.jecp.2012.09.014
- Kuo, L., & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist*, 41, 161-180.  
doi:10.1207/s15326985ep4103\_3
- Lindamood, P., & Lindamood, P. (1998). *The Lindamood Phoneme Sequencing Program for reading, spelling, and speech: LiPS: Teacher's manual for the classroom and clinic* (3<sup>rd</sup> ed.) Austin, TX: PRO-ED.
- Lovett, M. W., Lacerenza, L., & Borden, S. L. (2000). Putting struggling readers on the PHAST

- track: A program to integrate phonological and strategy-based remedial reading instruction and maximize outcomes. *Journal of Learning Disabilities*, 33, 458-476.  
doi:10.1177/002221940003300507
- Manis, F. R. (1985). Acquisition of word identification skills in normal and disabled readers. *Journal of Educational Psychology*, 77, 78-90
- McBride-Chang, C., & Kail, R. V. (2002). Cross-cultural similarities in the predictors of reading acquisition. *Child Development*, 73, 1392-1407.
- Mitchell, A. M., & Brady, S. A. (2014). Assessing affix knowledge using both pseudoword and real-word measures. *Topics in Language Disorders*, 34, 210-227.  
doi:10.1097/TLD.0000000000000020
- Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental Psychology*, 40, 665-681. doi:10.1037/0012-1649.40.5.665
- Nagy, W., Berninger, V., Abbott, R., Vaughan, K., & Vermeulen, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second-grade readers and at-risk fourth-grade writers. *Journal of Educational Psychology*, 95, 730-742.  
doi:10.1037/0022-0663.95.4.730
- Nagy, W., Carlisle, J. F., & Goodwin, A. P. (2014). Morphological knowledge and literacy acquisition. *Journal of Learning Disabilities*, 47, 3-12. doi:10.1177/0022219413509967
- Nation, K., & Cocksey, J. (2009). The relationship between knowing a word and reading it aloud in children's word reading development. *Journal of experimental child psychology*, 103, 296-308.

- Olson, R. K., Kliegl, R., Davidson, B. J., & Foltz, G. (1985). Individual and developmental differences in reading disability. In T. Waller (Ed.), *Reading research: Advances in theory and practice* (Vol. 4, pp. 1-64). London, England: Academic Press.
- Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing: An Interdisciplinary Journal*, 23, 189-208.
- Ouellette, G., & Fraser, J. R. (2009). What exactly is a yait anyway: the role of semantics in orthographic learning. *Journal of Experimental Child Psychology*, 104, 239-251.  
doi:10.1016/j.jecp.2009.05.001
- Perfetti, C. (1992). The representation problem in reading acquisition. In P. B. Gough & R. Treiman (Eds.), *Reading acquisition* (pp. 145-174). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Perfetti, C. A. (2003). The universal grammar of reading. *Scientific Studies of Reading*, 7, 3-24.  
doi:10.1207/S1532799XSSR0701\_02
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11, 357-383.
- Perfetti, C., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly*, 33, 283-319.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56-115. doi:10.1037/0033-295X.103.1.56
- Powell, M. J. (2009). The BOBYQA algorithm for bound constrained optimization without

- derivatives. Cambridge NA Report NA2009/06.
- Psychology Software Tools. (2014). E-Prime 2.0. Pittsburgh, PA: Author. Retrieved from <http://www.pstnet.com>
- Raaijmakers, J. G., Schrijnemakers, J. M., & Gremmen, F. (1999). How to deal with “the language-as-fixed-effect fallacy”: Common misconceptions and alternative solutions. *Journal of Memory and language*, 41, 416-426. doi:10.1016/j.jml.2005.12.010
- Ramirez, G., Chen, X., Geva, E., & Luo, Y. (2011). Morphological awareness and word reading in English language learners: Evidence from Spanish- and Chinese-speaking children. *Applied Psycholinguistics*, 32, 601-618. doi:10.1017/S0142716411000233
- Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *The Quarterly Journal of Experimental Psychology*, 61, 129-141. doi:10.1080/17470210701508822
- Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 482-503. doi:10.1037/0096-1523.25.2.482
- Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language*, 42, 342-364. doi:10.1006/jmla.1999.2687
- Rau, A. K., Moll, K., Snowling, M. J., & Landerl, K. (2015). Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *Journal of Experimental Child Psychology*, 130, 92–105. doi:10.1016/j.jecp.2014.09.012
- Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.



- Ricketts, J., Davies, R., Masterson, J., Stuart, M., & Duff, F. J. (2016). Preparation and delivery of three presentations for Culpeper teachers and administrators. *Journal of Experimental Child Psychology*, 150, 330-345.
- Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, 11, 235-257.  
doi:10.1080/10888430701344306
- Rueckl, J. G. (2016). Toward a theory of variation in the organization of the word reading system. *Scientific Studies of Reading*, 20, 86–97. doi:10.1080/10888438.2015.1103741
- Rueckl, J. G., & Raveh, M. (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain and Language*, 68, 110-117.  
doi:10.1006/brln.1999.2106
- Savage, R., Georgiou, G., Parrila, R., & Maiorino, K. (2018). Preventative reading interventions teaching direct mapping of graphemes in texts and set-for-variability aid at-risk learners. *Scientific Studies of Reading*, 22, 225-247. doi:10.1080/10888438.2018.1427753
- Schmalz, X., Marinus, E., & Castles, A. (2013). Phonological decoding or direct access? Regularity effects in lexical decisions of Grade 3 and 4 children. *Quarterly Journal of Experimental Psychology*, 66, 338-346. doi:10.1080/17470218.2012.711843
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523.
- Seymour, P. H. K., Aro, M., Erskine, J. M. and collaboration with COST Action A8 network (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94: 143–174. doi:10.1348/000712603321661859
- Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading

- acquisition. *Cognition*, 55, 151-218.
- Spencer, M., Muse, A., Wagner, R. K., Foorman, B., Petscher, Y., Schatschneider, C., & ... Bishop, M. D. (2015). Examining the underlying dimensions of morphological awareness and vocabulary knowledge. *Reading and Writing*, 28, 959-988. doi:10.1007/s11145-015-9557-0
- Stevens, M., & Brysbaert, M. (n.d.). *When do we have enough power in language research? Evidence from priming studies*. Unpublished manuscript, Department of Experimental Psychology, Ghent University, Ghent, Belgium. Retrieved from <http://crr.ugent.be/papers/When%20do%20we%20have%20enough%20power%20in%20language%20research.pdf>
- Taylor, J. S. H., Duff, F. J., Woollams, A. M., Monaghan, P., & Ricketts, J. (2015). How word meaning influences word reading. *Current Directions in Psychological Science*, 24, 322-328. doi:10.1177/0963721415574980
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2014). *Test of Word Reading Efficiency*. Austin, TX: Pro-Ed.
- Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, 124, 107-136. doi:10.1037/0096-3445.124.2.107
- Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language*, 29, 66-85. doi:10.1016/0749-596X(90)90010-W
- Tunmer, W. E., & Chapman, J. W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading*, 16, 122-140. doi:10.1080/10888438.2010.542527

- Van den Noortgate, W., De Boeck, P., & Meulders, M. (2003). Cross-classification multilevel logistic models in psychometrics. *Journal of Educational and Behavioral Statistics*, 28, 369-386. doi:10.3102/10769986028004369
- Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1994). Components of reading ability: Issues and problems in operationalizing word identification, phonological coding, and orthographic coding. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 279-332). Baltimore, MD: Paul H Brookes Publishing.
- Wagner, R. K., & Barker, T. A. (1994). The development of orthographic processing ability. In V. W. Berninger & V. W. Berninger (Eds.), *The varieties of orthographic knowledge, 1: Theoretical and developmental issues* (pp. 243-276). New York, NY, US: Kluwer Academic/Plenum Publishers.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing*. Austin, TX: Pro-Ed.
- Wang, H. C., Nickels, L., Nation, K., & Castles, A. (2013). Predictors of orthographic learning of regular and irregular words. *Scientific Studies of Reading*, 17, 369–384.
- Waters, G. S., Seidenberg, M. S., & Bruck, M. (1984). Children's and adults' use of spelling-sound information in three reading tasks. *Memory & Cognition*, 12, 293-305. doi:10.3758/BF03197678
- Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, 60, 502-529. doi:10.1016/j.jml.2009.02.001
- Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The Educator's Word Frequency Guide*. Touchstone Applied Science Associates.

Table 1  
*Demographics (N = 95)*

Variable	TA <sup>a</sup> (n = 49)		RD <sup>b</sup> (n = 46)	
	<i>M</i>	(SD)	<i>M</i>	(SD)
Age	9.75	(0.61)	9.63	(0.65)
TOWRE Phonemic Decoding Efficiency	110.35	(11.26)	80.75	(6.99)
TOWRE Sight Word Efficiency	109.18	(11.21)	82.24	(7.55)
	<i>n</i>	%	<i>n</i>	%
Grade				
Third	21	43	28	61
Fourth	28	57	18	30
Gender				
Female	24	49	29	63
Male	25	51	17	37
Received FRL				
No	43	88	29	63
Yes	4	8	10	22
NR	2	4	7	15
English Language Learner				
No	46	94	41	89
Yes	1	2	5	11
NR	2	4	0	0
Race				
African American	4	8	4	9
Asian/Indian	6	12	2	4
Hispanic	1	2	3	7
Caucasian/White	33	67	37	80
Multiethnic	3	6	0	0
Other	2	4	0	0
Has IEP				
No	45	92	28	61
Yes	4	8	15	33
Being evaluated	0	0	3	7
Has been retained				
No	49	100	43	93
Yes	0	0	3	7

*Note.* FRL = Free or reduced-price lunch; IEP = Individualized education plan; NR = Not reported; RD = Reading difficulty; TA = Typical achievement; TOWRE-2 = Test of Word Reading Efficiency-Second Edition (Torgesen et al., 2014). Percentages may not sum to 100 due to rounding.

<sup>a</sup> Criterion for TA was a TOWRE composite standard score > 35<sup>th</sup> %ile

<sup>b</sup> Criterion for RD was a TOWRE composite standard score < 25<sup>th</sup> %ile

Table 2

*Descriptive Statistics and Correlations for Analyzed Variables*

Descriptive Statistics and Correlations for Analyzed Variables					Correlations Among Variables										
Variable	Scale	Mean	(SD)	1	2	3	4	5	6	7	8	9	11		
Person x Item (Word) Variables ( <i>N</i> = 4,393)															
1	PSWR	Corr./Incorr. answer	.68	.47	—										
2	SK	Proportion correct	.78	.31	.18	—									
3	PSOC	Corr./Incorr. answer	.73	.44	.27	.19	—								
4	PK	Corr./Incorr. answer	.78	.41	.19	.11	.13	—							
Person (child) Variables ( <i>N</i> = 95)															
5	PPVT-4	Total correct	155.03	18.30	.18	.18	.16	.12	—						
6	OCT	Total correct	55.09	6.60	.24	.13	.30	.11	.42	—					
7	CTOPP E	Total correct	14.29	4.46	.24	.13	.16	.09	.36	.49	—				
8	CTOPP NR	Total correct	9.45	2.14	.17	.13	.14	.09	.31	.36	.39	—			
9	CTOPP RAN	Time elapsed	37.00	8.37	-.16	-.07	-.15	-.06	-.26	-.50	-.39	-.28	—		
10	TMS-D	Total correct	15.05	4.70	.24	.19	.19	.16	.71	.53	.43	.39	-.30		
Item (Word) Variables ( <i>N</i> = 48)															
11	Frequency <sup>a</sup>	Scaled count	51.88	7.39	.10	.08	.08	.11					—		
12	Consistency <sup>a</sup>	Proportion	.60	.39	.23	.17	.24	.19					-.04		

*Note.* Corr./Incorr. answer = Correct or incorrect answer, 1 for correct, 0 for incorrect; PSWR = Polysyllabic word reading; SK = Semantic knowledge; PSOC = Project-specific orthographic choice; PK = Phonological knowledge; PPVT-4 = Peabody Picture Vocabulary Test-Fourth Edition (Dunn & Dunn, 2007); OC = Orthographic Choice Test (Olson, Kliegl, Davidson, & Foltz, 1985); CTOPP = Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999); CTOPP E = CTOPP Elision; CTOPP NR = CTOPP Nonword Repetition; CTOPP RAN = CTOPP mean Rapid Digit and Rapid Letter Naming raw scores; TMS-D = Test of Morphological Structure Derivation subtest (Carlisle, 2000); Frequency = Standard Frequency Index from Zeno et al. (1995); Consistency = Stressed-syllable feedforward body-rime type-based consistency

<sup>a</sup>The correlations between frequency and consistency for PSWR, SK, PSOC, and PK are given because performance on item-specific measures varies by item. The correlations between the item and person variables are not given because they are null values; item characteristics do not vary across person and vice versa.

Table 3  
*Number of Matching PSWR, PK, PSOC, and SK Answers*

	PSWR		Total	PSWR and Other Match <sup>a</sup>
	Incorrect	Correct		
PK				
Incorrect	472	482	954	10%
Correct	950	2,489	3,439	57%
Total	1,422	2,971	4,393	67%
PSOC				
Incorrect	631	550	1,181	14%
Correct	791	2,421	3,212	55%
Total	1,422	2,971	4,393	69%
SK <sup>b</sup>				
Both incorrect	138	136	274	3%
One correct	582	811	1,393	13% <sup>c</sup>
Both correct	702	2,024	2,726	46%
Total	1,422	2,971	4,393	62% <sup>c,d</sup>

*Note:* PSWR = Polysyllabic word reading; PK = Phonological knowledge; PSOC = Project-specific orthographic choice; SK = Semantic knowledge task.

<sup>a</sup>Match is when children get both items correct or both items incorrect. <sup>b</sup>There were two semantic knowledge items for each word with yes/no answers for each. <sup>c</sup>This counts “one correct” as fitting the “incorrect” category, following the procedure we used for analysis. <sup>d</sup>The percentage of exact matches (excluding “one correct”) is 49%.

Table 4

*Model Results*

Fixed Effect	Model 1			Model 2			Model 3A			Model 3B		
	Word-specific			General			Word-specific + General			Model 3A + Morph. Aw.		
	<i>Coeff.</i> <sup>a</sup>	<i>(SE)</i>	<i>Prob.</i>	<i>Coeff.</i> <sup>a</sup>	<i>(SE)</i>	<i>Prob.</i>	<i>Coeff.</i> <sup>a</sup>	<i>(SE)</i>	<i>Prob.</i>	<i>Coeff.</i> <sup>a</sup>	<i>(SE)</i>	<i>Prob.</i>
$\lambda_0$ Intercept (Int.)	-2.405 <sup>••</sup>	-0.856	.08	-0.083 <sup>•</sup>	-0.774	.48	-0.912	-0.784	.29	-0.268	-0.78	.43
$\lambda_1$ Semantic Knowledge <sup>b</sup>	0.402 <sup>••</sup>	-0.147	.12				0.371 <sup>•</sup>	-0.15	.37	0.351 <sup>•</sup>	-0.15	.52
$\lambda_2$ Orthographic Knowledge <sup>b</sup>	0.456 <sup>•••</sup>	-0.099	.12				0.402 <sup>•••</sup>	-0.102	.38	0.399 <sup>•••</sup>	-0.102	.53
$\lambda_3$ Phonological Knowledge <sup>b</sup>	0.410 <sup>•••</sup>	-0.101	.12				0.380 <sup>•••</sup>	-0.102	.37	0.368 <sup>•••</sup>	-0.102	.52
Child <sup>c</sup>												
$\gamma_{010}$ TA <sup>b</sup>	1.114 <sup>•</sup>	-0.194	.22	0.526 <sup>•</sup>	-0.258	.61	0.533 <sup>•</sup>	-0.254	.41	0.347	-0.248	.52
$\gamma_{020}$ Grade 4 <sup>b</sup>	0.543	-0.543	.13	0.230	-0.22	.54	0.215	-0.22	.33	0.051	-0.216	.45
$\gamma_{030}$ PPVT-4 <sup>c</sup>				0.245 <sup>•</sup>	-0.101	.42-.54	0.212 <sup>•</sup>	-0.1	.25-.33	0.036	-0.112	.42-.44
$\gamma_{040}$ OCT <sup>c</sup>				0.297 <sup>•</sup>	-0.119	.41-.55	0.233 <sup>•</sup>	-0.118	.24-.34	0.184	-0.115	.39-.48
$\gamma_{050}$ CTOPP E <sup>c</sup>				0.253 <sup>•</sup>	-0.103	.42-.54	0.244 <sup>•</sup>	-0.102	.24-.34	0.271 <sup>••</sup>	-0.096	.37-.50
$\gamma_{060}$ CTOPP NR <sup>c</sup>				0.152	-0.085	.44-.52	0.128	-0.084	.26-.31	0.092	-0.080	.41-.46
$\gamma_{070}$ CTOPP RN <sup>c</sup>				0.058	-0.107	.46-.49	0.048	-0.106	.28-.30	0.004	-0.104	.43-.43
$\gamma_{080}$ TMS-D <sup>c</sup>										0.385 <sup>••</sup>	-0.123	.34-.53
Word												
$\gamma_{001}$ Frequency <sup>c</sup>	0.467 <sup>•••</sup>	(0.126) <sup>•••</sup>	.13	0.590 <sup>•••</sup>	-0.13	.62	0.488 <sup>•••</sup>	-0.123	.40	0.490 <sup>•••</sup>	-0.123	.55
$\gamma_{002}$ Consistency <sup>c</sup>	0.245 <sup>•</sup>	(0.111) <sup>•</sup>	.10	0.302 <sup>•</sup>	-0.121	.55	0.267 <sup>•</sup>	-0.112	.34	0.268 <sup>•</sup>	-0.112	.50
$\gamma_{003}$ Freq. x Cons.	-0.179	-0.115	.13	-0.179	-0.128	.65	-0.179	-0.118	.42	-0.179	-0.118	.58

Table 4 Continued

Random Effect	<i>SD</i>	Corr.	<i>SD</i>	Corr.	<i>SD</i>	Corr.	<i>SD</i>	Corr.
$r_{000}$ Child	0.627		0.566		0.554		0.515	
$r_{001}$ Frequency	0.272		0.192		0.183		0.182	
$v_{000}$ Word	1.424	-.45 <sub>TA</sub>	1.172	-.09 <sub>TA</sub>	1.061	.24 <sub>TA</sub>	1.068	.24 <sub>TA</sub>
		-.81 <sub>Grade</sub>		-.73 <sub>Grade</sub>		-.78 <sub>Grade</sub>		-.79 <sub>Grade</sub>
$v_{010}$ TA	0.484	-.45 <sub>Int.</sub>	0.365	-.09 <sub>Int.</sub>	0.349	-.24 <sub>Int.</sub>	0.348	-.24 <sub>Int.</sub>
$v_{020}$ Grade 4	0.487	-.81 <sub>Int.</sub>	0.45	-.73 <sub>Int.</sub>	0.454	-.78 <sub>Int.</sub>	0.456	-.79 <sub>Int.</sub>
$v_{030}$ PPVT-4			0.271		0.264		0.265	
$v_{040}$ Orthographic Choice <sup>†</sup>			0.37		0.365		0.366	
$u_{0000}$ Classroom	0.402		0.255		0.261		0.198	
$\omega_{00000}$ School	0.000		0.129		0.135		0.253	

*Note:* TA = Typical achievement; PPVT-4 = Peabody Picture Vocabulary Test-Fourth Edition (Dunn & Dunn, 2007); CTOPP E = Comprehensive Test of Phonological Processing, Elision subtest (Wagner, Torgesen, & Rashotte, 1999); CTOPP NR = Comprehensive Test of Phonological Processing, Nonword Repetition subtest (Wagner, Torgesen, & Rashotte, 1999); CTOPP RN = Comprehensive Test of Phonological Processing, Rapid Naming subtest (Wagner, Torgesen, & Rashotte, 1999); OCT = Orthographic Choice Test (Olson, Kliegl, Davidson & Foltz, 1985); TMS-D Test of Morphological Structure, Derivation subtest (Carlisle, 2000).

<sup>a</sup>Coefficients are given in log-odds. See Table 5 for information about interpreting log-odds in probability terms. The linear log-odds do not represent linear probabilities because of the 0 to 1 limits of the probability scale. <sup>b</sup>This is a dichotomous variable, so the coefficient represents the effect of getting the word right on the log-odds of a correct word reading response. The associated probability is for the case where the child has a correct score, compared to the intercept probability. <sup>c</sup>These variables have been standardized, so a 1-unit change represents the effect of a 1 *SD* greater score on that measure on the log-odds of a correct response. The associated probability ranges are for scores 1 *SD* below and 1 *SD* above the mean for the given variable, compared to the intercept probability such that all other scores are set to zero (dichotomous) or the mean (continuous).

▪  $p < .05$ . ▪▪  $p < .01$ . ▪▪▪  $p < .001$ .  $p$  values were calculated using nested chi-square difference tests in which models with and without each variable were estimated, and test was applied for the difference in deviance between the models.



Table 5  
*Relationship between Log-odds and Probabilities*

Log-odds Absolute Value	Probability	
	Positive Log-odds (Column A $\times$ +1)	Negative Log-odds (Column A $\times$ -1)
0.00	.50	.50
0.25	.56	.44
0.50	.62	.38
0.75	.68	.32
1.00	.73	.27
1.25	.78	.22
1.50	.82	.18
1.75	.85	.15
2.00	.88	.12
2.25	.90	.10
2.50	.92	.08
2.75	.94	.06
3.00	.95	.05

Figure 1

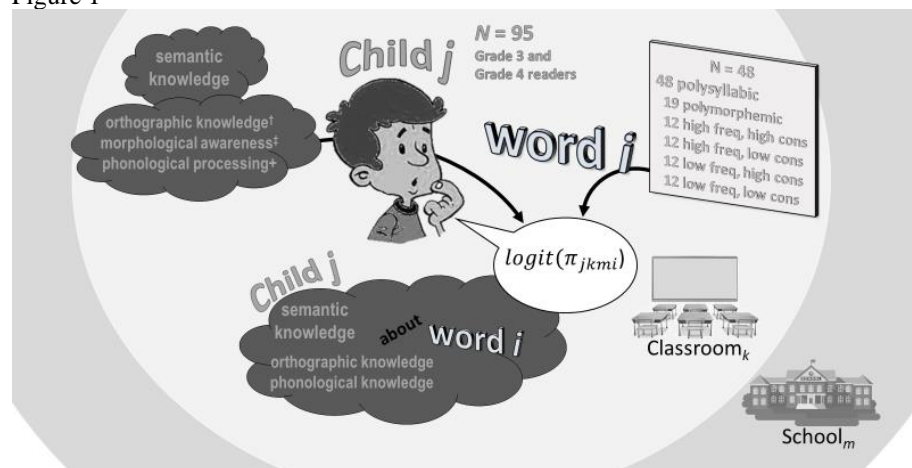


Figure 1B

$$\text{logit}(\pi_{jkmi}) = \lambda_0 + \gamma_{030} \text{PPVT4}_{jkm} + \gamma_{001} \text{Freq}_i + r_i + v_j + u_k + \omega_m$$

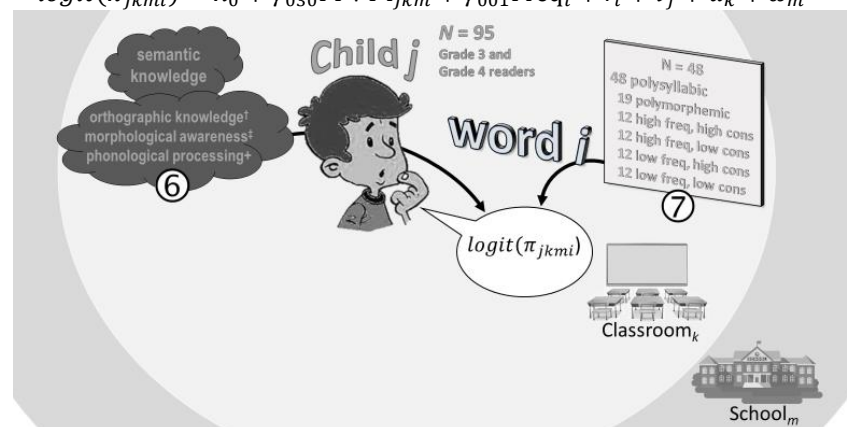


Figure 1A

$$\text{logit}(\pi_{jkmi}) = \lambda_0 + r_i + v_j + u_k + \omega_m$$

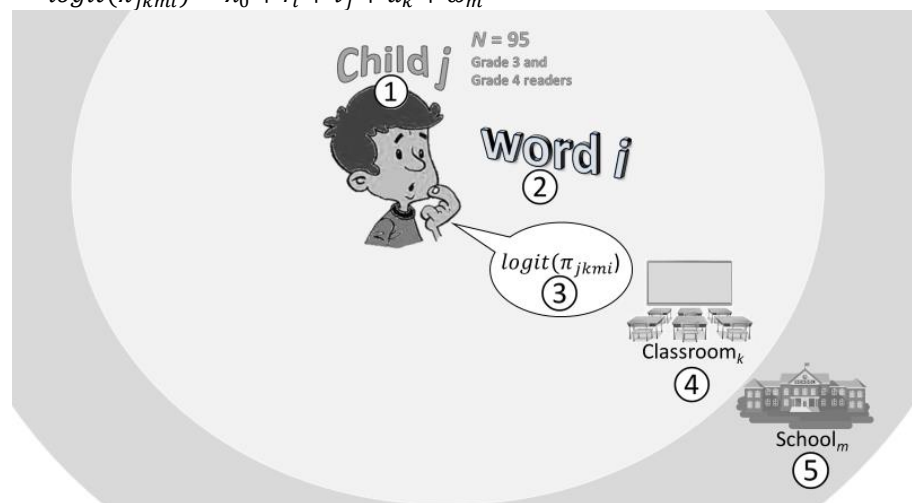
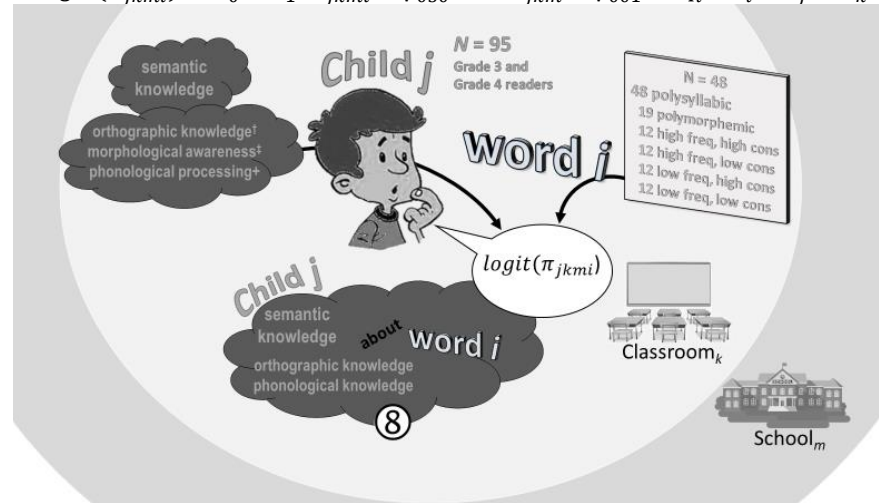


Figure 1C

$$\text{logit}(\pi_{jkmi}) = \lambda_0 + \lambda_1 \text{SK}_{jkmi} + \gamma_{030} \text{PPVT4}_{jkm} + \gamma_{001} \text{Freq}_i + r_i + v_j + u_k + \omega_m$$



*Figure 1.* A visual depiction of the mixed-effects model (also called generalized explanatory item-response model) representing this equation:  $\text{logit}(\pi_{jkmi}) = \lambda_0 + \lambda_1 \text{SK}_{jkmi} + \gamma_{030} \text{PPVT4}_{jkm} + \gamma_{001} \text{Freq}_i + r_i + v_j + u_k + \omega_m$ .  $\text{logit}(\pi_{ji})$  is a transformation of the outcome, the probability of a correct response. The transformation is required because probabilities—bounded by 0 and 1—have a nonlinear cumulative distribution function, the S-shaped curve and—if the outcome is not transformed, the effect of a change in an independent variable is greater in the middle of the distribution than at the ends. However, if the probability is transformed, linear regression techniques can be used and correct estimates of effects can be obtained. The transformation takes the predicted probability of a correct response,  $\pi_{jkmi}$ , and uses this equation:  $\text{logit}(\pi_{ji}) = \log\left(\frac{\pi_{jkmi}}{1-\pi_{jkmi}}\right)$  to express the probability in logits, a continuous metric. In the text and tables, coefficients for the independent variables are given for this continuous scale and are referred to as *log-odds*. Table 5 provides a guide for interpreting log-odds in terms of probabilities and illustrates the non-linear relationship between log-odds and probabilities.

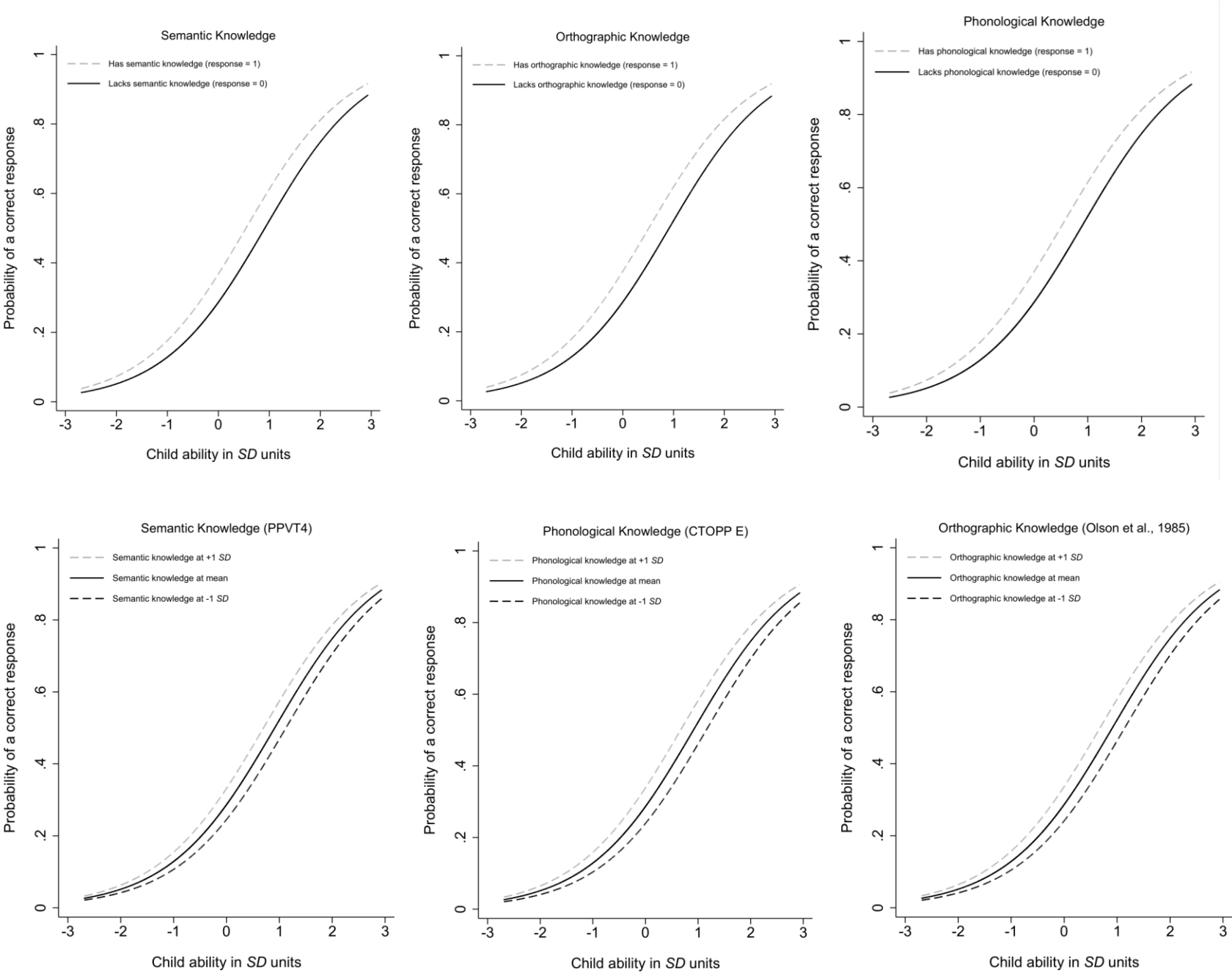
*Figure 1A.* Null random-effects model. The goal of this model to explain child  $j$ 's ① performance (correct or incorrect) when reading word  $i$  ②. The probability that child  $j$ 's answer is correct is  $\pi_{jkmi}$ , transformed to the logistic scale as  $\text{logit}(\pi_{jkmi})$  ③. Child  $j$  is a student in classroom  $k$  ④ and attends school  $m$  ⑤. In the equation for the null model without predictors,  $\text{logit}(\pi_{jkmi}) = \lambda_0 + r_i + v_j + u_k + \omega_m$ , there are four random effects that represent the variability in the outcome not explained by the independent variables. In this null model, there are no independent variables, so all of the variance is unexplained.  $r_i$  is the unexplained variability related to performance on word  $i$  for child  $j$  after the effects of the independent variables are included.  $v_j$  is a child random effect, the

unexplained variability related to child  $j$ 's performance reading word  $i$ .  $u_k$  is a classroom random effect, representing the unexplained variability related to the classroom child  $j$  is in, and  $\omega_m$  is a school random effect.

*Figure 1B.* Illustration of model for general child knowledge and word characteristics. Independent variables related to the specific child  $j$  and the specific word  $i$  are added here, represented with this equation,  $\text{logit}(\pi_{jkmi}) = \lambda_0 + \gamma_{110}\text{PPVT4}_{jkm} + \gamma_{101}\text{Freq}_i + r_i + v_j + u_k + \omega_m$ . In terms of the child variables, these variables—in the thought bubbles ⑥—represent child  $j$ 's *general child knowledge*, specifically, semantic, orthographic, phonological, and morphological knowledge and skills. In this example, the child's general semantic knowledge as measured by the PPVT4, for child  $j$ , is represented as  $\text{PPVT4}_{jkm}$ . In terms of the word variables, these are characteristics of the words—as shown on the board beside the word ⑦—that affect child performance, e.g., frequency,  $\text{Freq}_i$ . The fact that both child skills and word characteristics can be measured in the same model is one reason mixed-effects models are helpful.

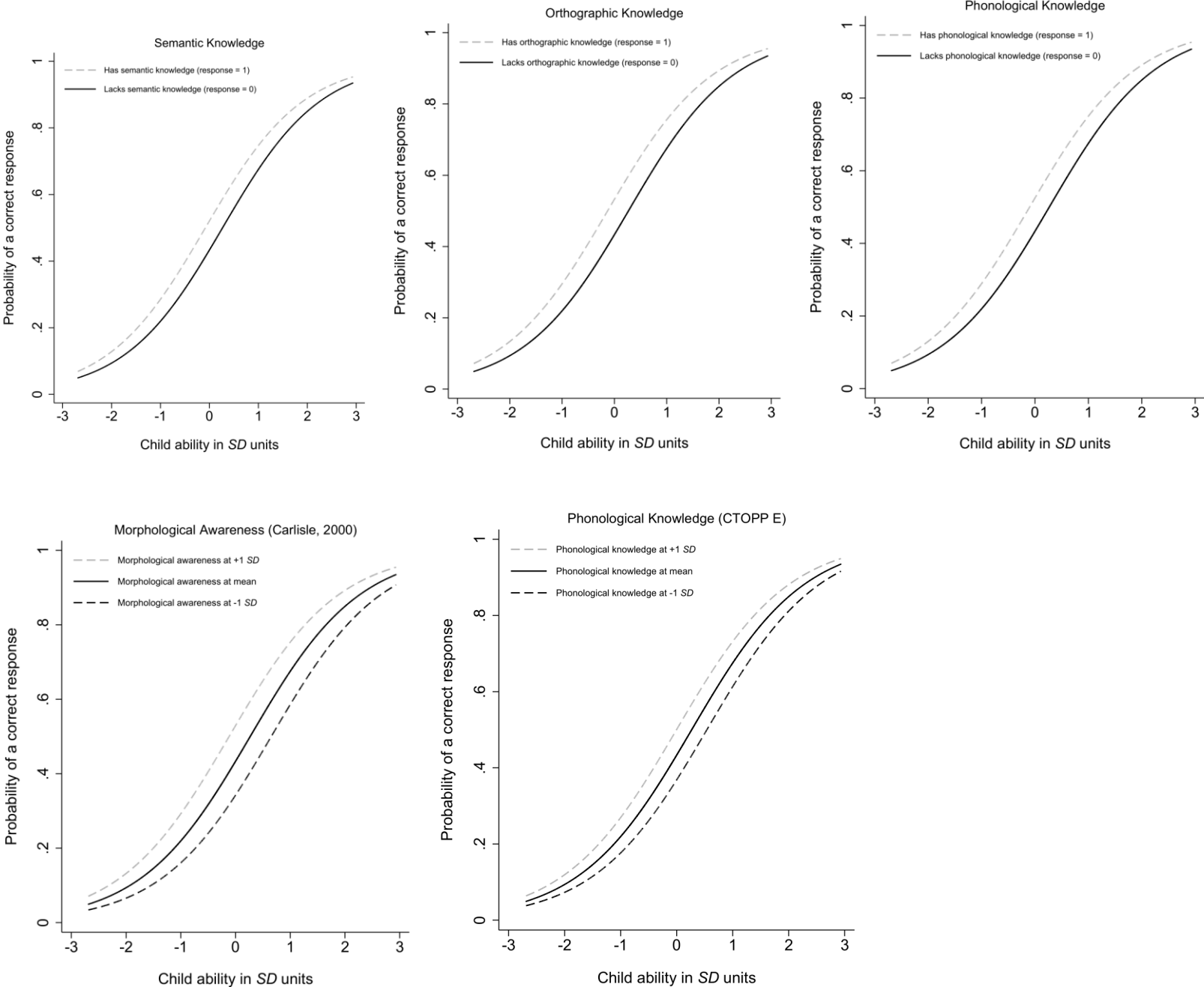
*Figure 1C.* Illustration of model for item-specific knowledge. The thought bubble ⑧ shows what Child  $j$  knows about Word  $i$  specifically, as given in this equation, which also includes the child and word independent variables:  $\text{logit}(\pi_{jkmi}) = \lambda_0 + \lambda_1\text{SK}_{jkmi} + \gamma_{030}\text{PPVT4}_{jkm} + \gamma_{001}\text{Freq}_i + r_i + v_j + u_k + \omega_m$ . In this equation,  $\text{SK}_{jkmi}$  represents the *item-specific child knowledge*, semantic knowledge child  $j$  (in classroom  $k$  and school  $m$ ) has about word  $i$  that might affect child  $j$ 's ability to read word  $i$  aloud (the outcome). The ability to include independent variables that measure the specific child's knowledge about the specific word is another reason mixed-effects models are unique and helpful. Copyright 2018. Used with permission.

Model 3A



*Figure 2.* For Model 3A, the probability of a correct response based on children's semantic, orthographic, and phonological knowledge of specific words (top section) and their general semantic, orthographic, and phonological knowledge if each of these was 1 *SD* above or below the sample mean (bottom section). For both general and specific knowledge, these figures represent these effects across a theoretical range of student ability. The lines follow the same logistic function because the models include no interactions. The interpretation of the bottom left graph is that children with high levels of semantic knowledge (PPVT4 Vocabulary score +1 *SD*; gray dashes) will have a higher likelihood of a correct response than children with average (black solid line) and low levels of semantic knowledge (-1 *SD*; black dashes). The distribution of child abilities reflects children's overall performance on the measures in our analyses. Children at -1 on the ability scale have overall performance 1 *SD* below the sample mean and children at +1 on the ability scale having overall performance 1 *SD* above the mean. These figures reflect for cases where all variables besides the variable of interest are held to their intercept values. This means that they are given for cases where a student had (1) sample-mean performance on all continuous measures and (2) incorrect responses for all word-specific tasks for (3) words of average difficulty for (4) a child with RD in third grade. Copyright 2018. Used with permission.

Model 3B



*Figure 3.* For Model 3B, the probability of a correct response based on word-specific (top section) and general knowledge (bottom section). General knowledge is plotted only for statistically-significant predictors, namely, morphological awareness (TMS-D) and phonological knowledge (CTOPP E) if each of these was 1 *SD* above or below the sample mean, top section). The interpretation is the same as for Figure 2. As in Figure 2, these figures reflect for cases where all variables besides the variable of interest are held to their intercept values. This means that they are given for cases where a student had (1) sample-mean performance on all continuous measures and (2) incorrect responses for all word-specific tasks for (3) words of average difficulty for (4) a child with RD in third grade. Copyright 2018. Used with permission.



## Appendix A

### Descriptive Statistics for Included Words ( $N = 48$ )

Appendix A, Table 1

*Descriptive Statistics for Included Words*

Word	SFI	Cons.	Word	SFI	Cons.
clarity	46.2	.28	resolve	46.9	1.00
correctly	53.6	1.00	respect	56.8	1.00
counsel	45.9	.41	return	60.8	1.00
country	65.9	.59	risky	45.5	1.00
manila	43.4	.10	seizure	44.8	.03
marriage	55.5	.28	sentence	61.4	.98
measure	59.1	.27	shortly	54.2	.98
member	58	1.00	similar	61	.86
memory	56.9	1.00	slightly	57.4	1.00
mileage	43.1	1.00	slowly	62.9	.42
motel	44.4	1.00	special	65	.01
moveable	47.1	.17	specify	46	.10
parental	46.8	.98	speedy	44.7	1.00
petition	45.3	.10	stature	44.7	.01
presence	56.6	.17	stirrup	43.1	.32
pretend	53	1.00	stomach	56.5	.11
pretty	59.4	.02	suspense	44.8	1.00
public	62.9	.97	sustain	45.2	.69
publish	46.5	.97	sweater	50.6	.09
purity	44.1	.06	symphony	43.9	1.00
religion	56.6	.10	tariff	46.4	.28
removal	49	.17	together	66.7	.79
renew	43.1	.84	torment	43.5	1.00
repair	53.8	.65	trouble	61.1	1.00

*Note:* SFI = Standard Frequency Index from Zeno et al. (1995).

Cons. = Type-based feedforward stressed syllable rime unit consistency

## Appendix B

### Word-Specific Phonological Knowledge Task Construction

Appendix B, Figure 1 shows an item on the phonological knowledge task. To generate syllables without prosody, a linguist recorded pronunciations of each syllable by saying them in succession but with a small amount of space between them. For example, *suspense* was pronounced this way: /sə/, 500 ms of silence, /spɛns/. We followed the syllable division principles proposed by Treiman and Zukowski (1990) based largely on Clements and Keyser (1983). Unstressed vowels retained their /ə/ or syllabified consonant sounds (e.g., final /l/ in *parental*) but stress was equal for each syllable. examiners recorded the accuracy of each response after each item. Independent scorers independently coded the responses recorded by the original test examiners, and interrater reliability was calculated; Cohen's  $\kappa$  was .91.

To further test the validity of this measure, we conducted an experiment with adults ( $N = 22$ ) in which they completed the phonological knowledge task with the syllables in reverse order (e.g., *renew* as /nu ... ri/). They also did the polysyllabic word reading task. We recorded their response times for each. Adults' phonological knowledge and polysyllabic word reading response times were highly correlated,  $r = .80$ , so we judged that this syllable task (albeit made more difficult for adults) captured a reading-related skill.

## Appendix C

### Word-Specific Semantic Knowledge Task Construction

For each of the 48 polysyllabic words, two sentences were constructed, one in which the sentence made sense based on the meaning of the polysyllabic word and one where it did not. For example, for the target word *suspense*, the correct sentence was “The ending of the movie left everyone in suspense.” and the incorrect sentence was “The delicious dessert made everyone full of suspense.” Appendix C, Figure 1 provides an example item pair.

The sentences were assigned to one of two lists. For each pair of sentences, one was randomly assigned to the first and one to the second. For each list, the examiner showed the student the items written on a piece of paper and also read them aloud. The test items for each list were preceded by three practice items to help the children understand what they were supposed to do. The children were asked to say whether the sentence made sense or not, by saying “makes sense” or “yes” or “doesn’t make sense” or “no.” Items were scored 1 for a correct response and 0 for an incorrect response. For data analysis, the average of the two items was taken, such that 1 meant the child gave a correct response for both items and 0 that the child gave an incorrect response for both items. This placed the word-specific semantic scores on the same scale as the orthographic and phonological ones.

One possible concern about the test was that child performance might be affected by the frequency of other words in the target sentences. To address, we counted the occurrences of each word in the sentences on the word-specific semantic knowledge task. There were 866 total non-target words, representing 388 unique words. Non-target words were neither any of the 48 target words nor proper nouns. The mean frequency for the types of non-target words from the 6.56 million word *EWFG* corpus for first to eighth grade was 10,678 per 6.56 million words.

The mean frequency of all words in the corpus was 383 per 6.56 million words.

We then assessed the risk that a word might be low enough in frequency to affect performance on the semantic knowledge measure, based on two possibilities. One option was that a word was below the mean frequency for all 6.56 million words, i.e., less than 383. The other was that a word would have a standard frequency index (SFI) below 55. Zeno et al. (1995) gave examples for words at different SFI cutpoints: 60 ~ “brother, minute, yellow”; 55 ~ “bowl, dozen, hang, snake”; 50 ~ “abuse, sphere, grammar, charter” (p. 13). We judged that words above 55 were likely to be known by most students. There were 17 words that met one or both criteria for being a lower-frequency non-target word, termed a *risk word*. We linked these 17 words to the targets by finding sentences that contained by a risk word and a target. For example, one sentence was “Our teacher had us write a sentence to practice handwriting.” It contained the risk word “handwriting” and the target word “sentence.”

We ran two regressions. In the first, the score on the semantic knowledge test was the outcome and risk-word status the only predictor. There was no effect,  $\Delta\chi^2_{[1]} = 0.004, p = .95$ . In the second, we included risk-word status in the word-level regression containing word-level effects. There was again no effect,  $\Delta\chi^2_{[1]} = 0.003, p = .96$ . We concluded there was no effect of the frequency of the risk words on the probability that a child would pronounce a given word correctly.

## Appendix D

### Using a Grapheme-Phoneme Consistency Measure

Because we thought consistency of grapheme-phoneme links could influence accuracy, possibly moderated by frequency, we chose words based on consistency rather than regularity, following the convention of prior reading studies involving polysyllabic stimuli (Chateau & Jared, 2003; Jared & Seidenberg, 1990; Yap & Balota, 2009). Although consistency has been measured in a variety of ways (see Yap & Balota, 2009 for multiple options), we examined the consistency of the relationship between the syllable body and the phonological rime, equivalent to the word body-rime approach used by Glushko (1979) in monosyllabic words. This was one of the methods used by Yap and Balota (2009) and for which they observed consistency effects. We used feedforward consistency, following Jared and Seidenberg (1990) and Chateau and Jared (2003), mapping from the letters (the syllable body) to sound (the rime). It is calculated by first counting the number of *friends* a word has. Friends are words with the same spelling of a given syllable body that also have the same pronunciation of the rime. Then, the total number of friends and *enemies* is calculated. Enemies are words with the same spelling for the syllable body but a different pronunciation. Feedforward consistency is then calculated by dividing friends by friends and enemies. The method of calculation for *pretty* is explained in Figure Appendix D, Figure 1.

### Procedure for Testing Validity of Chosen Consistency Measure

To choose among several possible consistency measures, we also ran exploratory analyses using lexical decision accuracy and RT data for all polysyllabic words in the English Lexicon Project (Balota et al., 2007), following the Yap and Balota (2009) seven-step process. We compared stressed syllable body/rime type consistency to the same metric with token

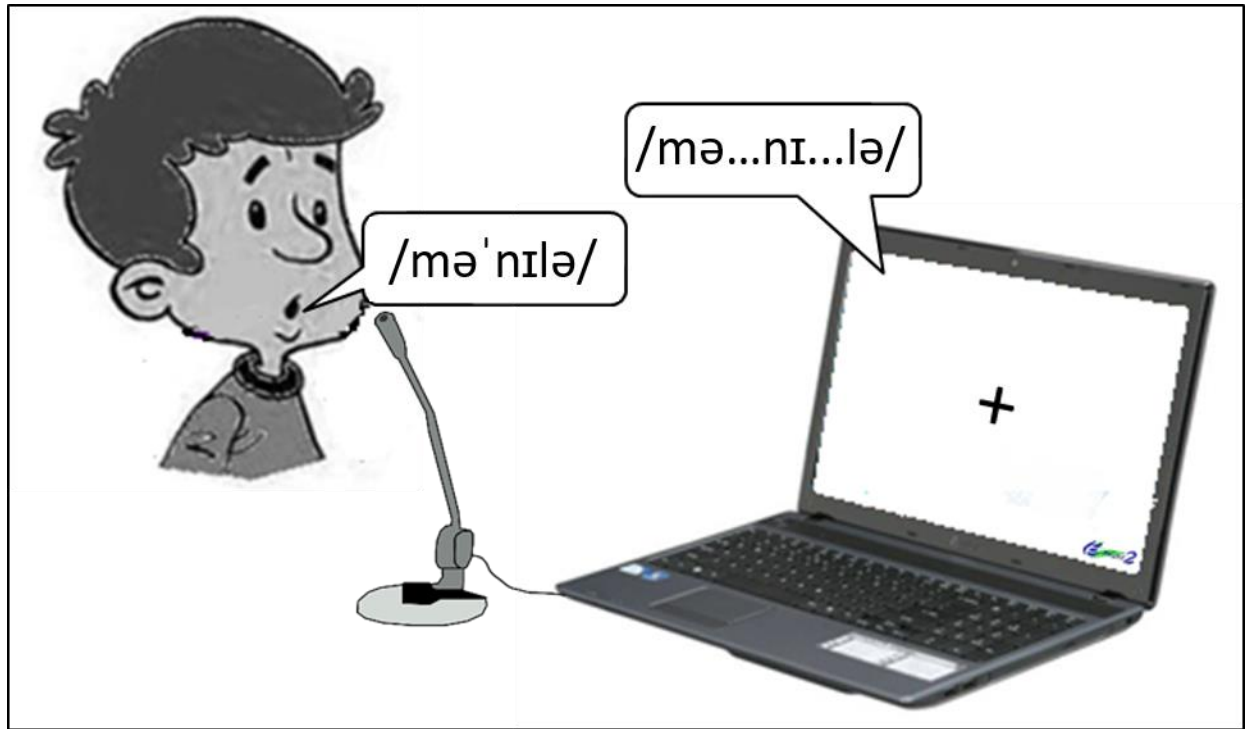
consistency, as well as to mean type and token consistency across syllables in a word and type and token consistency for the first syllable in the word. *Type-based stressed syllable consistency* was one of the strongest measures, so we chose this metric. From a theoretical point of view, Arciuli, Monaghan, and Ševa (2010) have shown that readers use words' orthographic cues to determine stress assignment, so the stressed syllable may already receive more attention from the reader. Moreover, calculating consistency across all syllables would have reduced the consistency of rime bodies. For example, *o* generally represents /oo/ (*meter*) or /ɑ/ (*comic*) but can be /ə/ (*reason*) in an unstressed syllable. Including unstressed syllables in the calculation would have reduced consistency without sensitivity to what readers may have intuited about stress placement.

Appendix D, Table 1

*Considerations in Choosing a Polysyllabic Consistency Measure*

Consideration	Options	Examples
Direction of mapping	<b>Letter-to-sound (feedforward)</b> Sound-to-letter (feedback)	<i>risky</i> : <i>isk</i> →/ɪsk/; <i>y</i> →/i/ <i>pretty</i> : /it/→/ɪt/; /i/→y
Syllabification chosen	<b>Orthographic syllable</b> Phonological syllable	<i>ab-ac-us</i> /'æ-bə-kəs/
Syllables evaluated	First syllable Syllable with primary stress <b>Syllables with pri. and sec. stress</b> All syllables (composite)	<i>petition</i> : <i>pe</i> <i>suspense</i> : <i>pence</i> <i>pronunciation</i> : <i>un, a</i> <i>renew</i> : <i>e, ew</i>
Unit of analysis	Onset <b>Syllable body/Rime</b> Syllable body/Rime + one letter Whole syllable Body of the BOSS (BOB) Basic orth. syll. structure (BOSS) Other (V1C1, V1C2, V1, V2, etc.)	<i>seizure</i> : <i>s, z</i> <i>sentence</i> : <i>en, ence</i> <i>suspense</i> : <i>usp, ence</i> <i>religion</i> : <i>re, li, gion</i> <i>respect</i> : <i>esp</i> <i>torment</i> : <i>torm</i>
Weighting by frequency	Token ( <i>n</i> words × word freq) <b>Type (<i>n</i> words with unit)</b>	<i>pretty friends</i> = 2133 <i>pretty friends</i> = 17

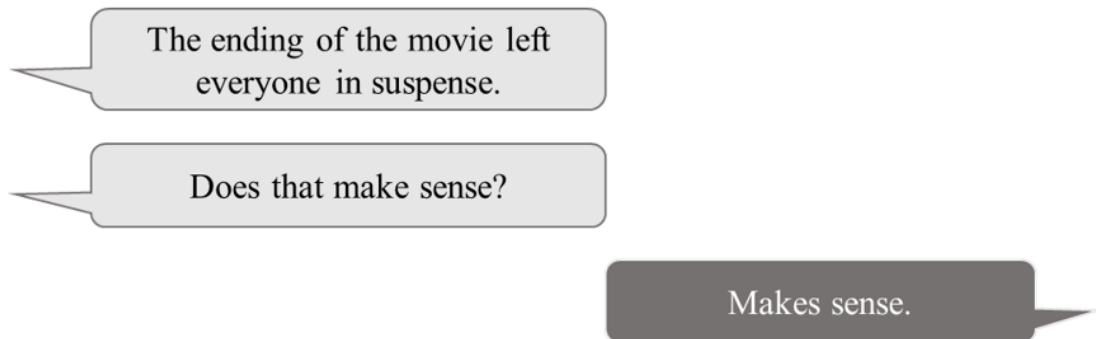
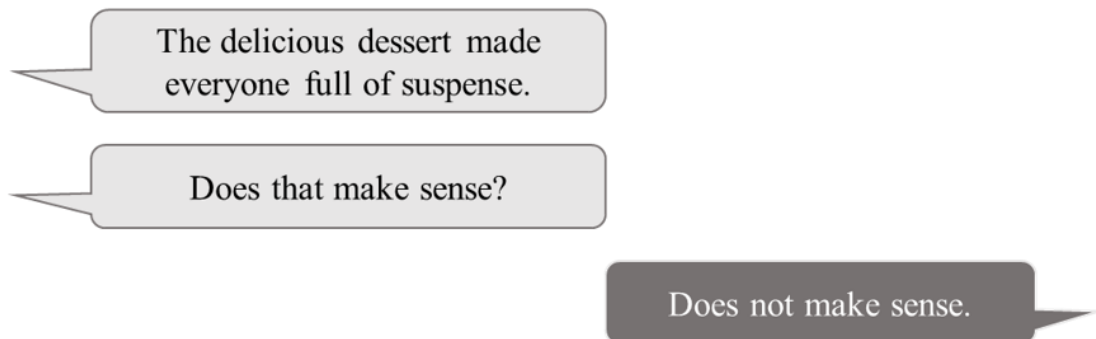
*Note:* The consistency choices made for this study are given in bold. The orthographic and phonological syllables were those given by CELEX. The former represents typical dictionary syllabification and the latter phonological syllable division based on CV phonology (Clements & Keyser, 1983; Treiman & Zukowski, 1990). The VC structures considered are based on the coding of Chateau and Jared (2003). The BOSS and BOB measures are described by (Taft, 1979) and (Taft, 1992), respectively. V = vowel; C = consonant.



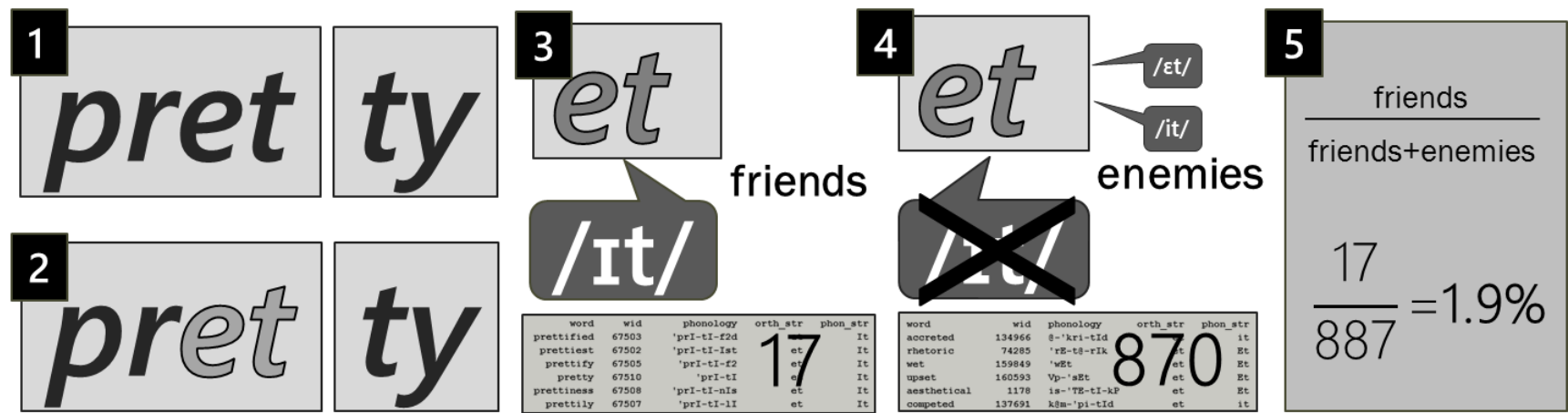
*Appendix B, Figure 1.* Illustration of administration procedure for phonological knowledge task.

This test measured children's word-specific phonological knowledge, that is, whether the child has sufficient segmental phonology for a given word to blend the word's verbally presented syllables and pronounce the correct word. The syllables were read without prosody and with 500 ms pause between each syllable (e.g., /mə/, -500 ms pause, /nɪ/, 500 ms pause, /lə/). The stressless syllables were recorded and presented using a computer to ensure that the variability on children's responses was not due to inconsistencies in item presentation. Children sat facing the computer screen and a microphone connected to the computer. They listened to the segmented pronunciation of each of the 48 target polysyllabic words, presented in a random order, then provided their response. The test administrator recorded the accuracy of their responses. Natural pronunciation of the words was required for correct responses. Copyright 2015. Used with permission.



**Semantic Knowledge Test  
Part 1****Semantic Knowledge Test  
Part 2**

*Appendix C, Figure 1.* Illustration of semantic knowledge task items and correct responses. Part 1 refers to the first half of the test and Part 2 to the second half, sections that were completed at different times. Copyright 2018. Used with permission.



Appendix D, Figure 1. Calculation of feedforward stressed syllable body-rime consistency coding procedure for pretty. [1] The word is divided into syllables using orthographic syllabification. [2] The syllable body is identified. [3] The syllable body’s friends (same rime pronunciation for that body) are identified. [4] The syllable body’s enemies (different rime pronunciation for that body) are identified. [5] Consistency is the proportion of words with the same syllable body that are friends. In this study, it is a type-based measure, meaning that each word is counted once, and words are not weighted by frequency. Copyright 2015. Used with permission