*2) Title Page (WITH Author Details)

Running head: SEMANTICS AND WORD-LEVEL READING

Evidence for semantic involvement in regular and exception word reading in emergent readers of English

Jessie Ricketts[1], Robert Davies[2], Jackie Masterson[3], Morag Stuart[3], & Fiona J. Duff[4]

[1]Royal Holloway, University of London

[2]University of Lancaster

[3]UCL Institute of Education

[4]University of Oxford

Correspondence:

Jessie Ricketts

Department of Psychology, Royal Holloway, University of London, Egham Hill, Egham, Surrey, TW20 0EX, UK

Email: jessie.ricketts@rhul.ac.uk

Telephone: (0044) 1784 414 623

Acknowledgments

We acknowledge the children, families and schools who participated; and Lizzie Penn, Natalie McConnachie and Dan Greene for their assistance. This research was funded by the Experimental Psychology Society, the Institute of Education (University of London) and the University of Reading. The first author is supported by the Economic and Social Research Council (grant number ES/K008064/1) and the last author by the Nuffield Foundation (grant number EDU/40062).

Running head: SEMANTICS AND WORD-LEVEL READING

Evidence for semantic involvement in regular and exception word reading in emergent readers of English

Abstract

We investigated the relationship between semantic knowledge and word reading. Twenty-seven six-year-old children read words both in isolation and in context. Lexical knowledge was assessed using general and item-specific tasks. General semantic knowledge was measured using standardised tasks in which children defined words and made judgements about the relationships between words. Item-specific knowledge of to-be-read words was assessed using auditory lexical decision (lexical phonology) and definitions (semantic) tasks. Regressions and mixed-effects models indicated a close relationship between semantic knowledge (but not lexical phonology) and both regular and exception word reading. Thus, in the early stages of learning to read, semantic knowledge may support word reading irrespective of regularity. Contextual support particularly benefitted reading of exception words. We found evidence that lexical-semantic knowledge and context make separable contributions to word reading.

Keywords: semantic, word reading, context, vocabulary, lexical phonology, mixed-effects models

Evidence for semantic involvement in regular and exception word reading in emergent readers of English

Knowledge of the meaning of words and phrases (semantic knowledge) has an important role to play in reading. Logically, a child needs to understand the meaning of the words and phrases contained within a text in order to fully understand it. The Simple View of Reading (e.g., Gough & Tunmer, 1986), an influential framework for understanding reading comprehension, posits that successful reading comprehension is underpinned by oral language comprehension (including semantic knowledge) as well as word reading abilities. Indeed, studies adopting longitudinal and experimental (randomised controlled trial) designs (e.g., Clarke, Snowling, Truelove, & Hulme, 2010; Nation & Snowling, 2004) have yielded convincing evidence that semantic knowledge is causally related to reading comprehension ability.

There is also evidence that oral language ability contributes to the development of word reading in children, with influences from both phonology and semantics (e.g., Duff & Hulme, 2012; Nation & Cocksey, 2009; Nation & Snowling, 2004; Ouellette & Beers, 2010; Ricketts, Nation, & Bishop, 2007). We concentrate here on semantic influences. Nation and Snowling (2004) showed that semantic knowledge at age 8 years predicted later word reading at age 13 years, after accounting for decoding ability, phonological skills and the autoregressor (word reading at 8 years). In an extension of this research, Ricketts et al. (2007) demonstrated a more specific relationship: that oral vocabulary knowledge was more closely associated with exception word reading than regular word reading. Exception words are words with unusual mappings between spelling and sound (e.g., <yacht>, <pint>) whereas regular words contain only predictable spelling-sound mappings. Importantly, regular words can be readily decoded using knowledge of the usual relationships between spelling patterns (graphemes) and sounds (phonemes) whereas exception (or irregular) words cannot (e.g., using such a strategy would result in <yacht> being pronounced to rhyme with “matched” rather than “cot”). Regular words are usually read more accurately than exception words by typically developing children (e.g., Nation & Cocksey, 2009).

In the literature outlined above, receptive and/or expressive oral vocabulary measures have typically been used to assess semantic knowledge. It is worth noting that the acquisition of oral vocabulary or lexical-semantic knowledge is incremental rather than an all-or-nothing process, with individuals adding to existing lexical-semantic representations, as well as acquiring new representations, throughout the lifespan. Studies conducted by Ouellette and colleagues (e.g., Ouellette, 2006; Ouellette & Beers, 2010) have acknowledged this by making a distinction between breadth (number of words known) and depth (what is known) in vocabulary knowledge. Ouellette and Beers found that for children aged 5–7 years their depth measure was a significant predictor of exception word reading whereas their breadth measure was not; the reverse pattern was observed for older readers (11–12 years).

Oral vocabulary is an important part of semantic knowledge. However, semantic knowledge additionally encompasses an understanding of the meaning-based relationships between words, the meaning of phrases and so on. As far as we have ascertained, the study by Nation and Snowling (2004) is unique in investigating the relationship between semantic knowledge and word reading by not only using the usual measure of oral vocabulary (in this case an expressive measure), but also a measure that goes beyond such lexical-semantic knowledge – a composite of ‘semantic skills’ comprising semantic fluency and synonym judgement. In regression analyses, Nation and Snowling found that their two measures of semantic knowledge made equivalent contributions to explaining variance in word reading, as measured concurrently and longitudinally by a well-established standardised test. However, their analysis of exception word reading, more specifically, showed that oral vocabulary at age 8 years was a significant predictor of exception word reading four years later, whereas the semantic composite was not.

A number of mechanistic accounts for the relationship between semantic knowledge and word reading have been proposed. Walley, Metsala, and Garlock, (2003) suggested that the relationship between semantic knowledge and word reading is indirect. According to their lexical restructuring hypothesis, oral vocabulary development serves to specify phonological representations, which in turn are critical for word reading development (e.g., Bishop & Snowling, 2004; Brady & Shankweiler, 1991; Goswami & Bryant, 1990). Computational models of word reading assume a more direct relationship. In the triangle model, words can be read aloud via two pathways, including one that maps indirectly from orthography to phonology via semantics (Harm & Seidenberg, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996). The dual route cascaded model (DRC; Coltheart, Rastle, Perry, Langdon & Ziegler, 2001) also makes reference to a semantic route; however, this route has not been implemented in its simulations, and the activation of semantics is not necessary for word reading. In the triangle model, semantic knowledge is necessary, and has a particularly important role to play in the reading of exception words, and for poor readers. Similarly, in his developmental account, Share (1995) has argued that top-down support from semantic information helps readers to resolve decoding ambiguity (for similar proposals, see Bowey & Rutherford, 2007; Tunmer & Chapman, 2012). According to this view, when a word is encountered that cannot be readily decoded, either because it is an exception word, or the reader does not possess the requisite reading ability, semantic information relating to the context or the word can be combined with a partial decoding attempt to successfully read the word.

In most studies, the relationship between semantic knowledge and word reading has been investigated by measuring both constructs and testing whether these constructs are correlated across participants, showing that there is a general relationship between some index of the semantic knowledge that individuals can access, and the number of words that they can read on an unrelated measure. However, theoretical positions proposing a direct and necessary relationship between semantics and word reading (e.g. Harm & Seidenberg, 2004) motivate a more precise hypothesis of the relationship between these variables: specifically, that knowledge of an individual word should aid reading of that particular word. This hypothesis is corroborated by evidence from semantic dementia patients, some of whom experience difficulty reading exception words alongside their semantic impairments, but who are more likely to successfully read exception words for which they know the meaning (Graham, Hodges, & Patterson, 1994; Woollams, Ralph, Plaut, & Patterson, 2007; but see Schwartz, Saffran & Marin, 1980, for a contrasting case). In what follows, we will summarise pertinent data from studies with children.

Nation and Cocksey (2009) probed item-level relationships between semantic knowledge and word reading in children. Participants aged 7 years read lists of regular and exception words and completed auditory lexical decision and definitions tasks as, respectively, indices of phonological and semantic lexical knowledge. Nation and Cocksey found that children demonstrated phonological and semantic knowledge of the majority of words that they read correctly, and this relationship was stronger with exception than regular words, though a small percentage of words were read correctly without being recognised in the auditory lexical decision task or defined correctly. Across-items performance in both auditory lexical decision and definitions tasks showed equivalent correlations with word reading. In further analyses, both auditory lexical decision performance and definitions knowledge were entered into by-items regression analyses predicting exception word reading. Auditory lexical decision performance explained unique variance in exception word reading after accounting for the variance explained by definitions performance. However, definitions did not explain unique variance in exception word reading after accounting for the variance explained by auditory lexical decision. This led the authors to conclude that lexical phonology (familiarity with a word's phonological form) is sufficient to support word reading, and that possessing deeper semantic knowledge does not predict more successful reading. However, they interpret their findings with caution due to the small sample size and the recognition that by-items performance on their auditory lexical decision task was skewed towards ceiling.

Two training studies conducted by Duff and Hulme (2012, Experiment 2) and McKague, Pratt, and Johnston (2001) have shown that pre-exposing children to the phonological forms of words facilitates learning to read those items, as does pre-exposure to phonology plus semantics (see also Ouellette & Fraser, 2009; Wang, Nickels, Nation, & Castles, 2013). In Duff and Hulme, and McKague et al., pre-exposure to phonology plus semantics does not confer an additional advantage beyond pre-exposure to phonology alone, resonating with Nation and Cocksey's (2009) claim that lexical phonology is sufficient to support word reading. In contrast, adult studies indicate that semantic pre-exposure supports learning to read exception words, over and above pre-exposure to phonology alone (McKay, Davis, Savage, & Castles, 2008; Taylor, Plunkett, & Nation, 2011), consistent with data from semantic dementia patients (for a review of relevant research, see Taylor, Duff, Woollams, Monaghan, & Ricketts, 2015). Taken together, findings are mixed. In relation to ideas put forward by Share (1995) and others (Bowey & Rutherford, 2007; Tunmer & Chapman, 2012), knowing a word's phonological form may be sufficient to support partial decoding attempts but knowledge of semantics may also be important. Resolving this issue was one motivation for our study.

The Present Study

We investigated whether semantic knowledge predicts word reading in 6–7 year-old children, bringing together two approaches that have been used to explore this relationship. In the first approach, we measured semantic knowledge and word reading using standardised tests and also asked children to read lists of regular and exception words to assess whether there is a general relationship between semantic knowledge and word reading (cf. Nation & Snowling, 2004; Ouellette & Beers, 2010; Ricketts et al., 2007). As in Nation and Snowling (2004), we measured both lexical-semantic knowledge (expressive vocabulary) and broader semantic knowledge (semantic relations between words). Our measure of lexical-semantic knowledge was an expressive oral vocabulary measure that captured depth as well as breadth; such measures have been found to predict exception word reading more strongly than measures of breadth alone in children of this age (Ouellette & Beers, 2010). Our measure of broader semantic knowledge assessed awareness of meaning-based relationships between words.

In our second approach, we investigated item-specific relationships between word knowledge and word reading (after Nation & Cocksey, 2009). We exposed children to lists of regular and exception words in tasks assessing word knowledge (auditory lexical decision, definitions) and reading (reading in isolation, reading in sentence context) in order to probe whether knowledge of a word's phonological form or semantic attributes would predict the ability to read that particular word. Our study builds on previous work by assessing reading in a more naturalistic contextualised task, in addition to the reading in isolation approach adopted by the majority of studies. Notably, children typically read words, particularly exception words, more accurately in context (Archer & Bryant, 2001; Nation & Snowling, 1998). We also extend previous research by using mixed-effects models to estimate item-specific relationships between word knowledge and word reading while accounting for error variance due to participants and items.

In sum, we took a novel approach to probing the mechanisms underpinning the relationship between word knowledge and word reading by: 1) investigating general and item-specific relationships in the same study with the same children, 2) measuring richer semantic knowledge using the semantic relationships task, as well as oral vocabulary, and 3) measuring word reading in context as well as in isolation. Our hypotheses were as follows. First, we hypothesised a general relationship between semantic knowledge (both vocabulary and semantic relationships) and word reading (Nation and Snowling, 2004), that would be stronger for exception than regular words (Ricketts et al., 2007). Second, we predicted an item-specific relationship between word knowledge (as indexed by auditory lexical decision and definitions) and word reading, again expecting that this relationship would be stronger for exception words (Nation & Cocksey, 2009). We further predicted that auditory lexical decision might be an equivalent or stronger predictor of word reading than definitions (Nation & Cocksey, 2009). Finally, we expected that regular words would be read more accurately than exception words (Nation & Cocksey, 2009), words would be read more accurately in context than in isolation (Archer & Bryant, 2001), and this contextual facilitation effect would be more pronounced for exception words than regular words (Nation & Snowling, 1998; Share, 1995).

Method

Participants

A sample of 27 children (10 boys) aged 6-7 years participated in this study (*M* = 6.50, *SD* = .26). All children aged 6-7 years attending two schools serving socially mixed catchment areas in Birmingham, UK were invited to take part provided they spoke English as a first language and did not have any recognised special educational need. Data were collected and analysed from all children for whom informed parental consent was received. Children had experienced two years of formal literacy instruction. Ethical approval was provided by the ethics committee at the Institute of Education, University of London.

Materials and procedure

Standardised tasks. Children completed standardised tasks in two sessions, each lasting approximately 30 minutes. Sessions were separated by approximately one week (*M* days between testing sessions = 5.26, *SD* = 1.58). All background measures were published standardised tasks and were administered according to manual instructions, in a fixed order across the two sessions.

Nonverbal reasoning was measured using the Matrix Reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (WASI, Wechsler, 1999), which is a pattern completion task.

Word-level reading was assessed using the Phonemic Decoding Efficiency (PDE) and Sight Word Efficiency (SWE) subtests of the Test of Word Reading Efficiency (TOWRE, Torgesen, Wagner, & Rashotte, 1999). In each subtest, children are asked to read a list of nonwords (PDE) or words (SWE) of increasing length and difficulty as quickly as they could. Efficiency is indexed by the number of nonwords or words read correctly in 45 seconds.

Semantic knowledge was indexed by the Vocabulary and Similarities subtests of the WASI (Wechsler, 1999). The Vocabulary subtest is a measure of expressive vocabulary that requires children to verbally define words. The Similarities subtest measures knowledge of the semantic relationships between words; children are presented with two semantically related words and are asked to describe how these words are related in meaning.

Experimental tasks. Children were exposed to 40 words in the context of four tasks, two assessing reading (reading in isolation, reading in context) and two indexing lexical knowledge (auditory lexical decision, definitions). Tasks were completed in the following fixed order: auditory lexical decision; reading in isolation; definitions; and reading in context. Tasks were presented in this order to limit contamination across tasks. Nonetheless, repetition effects were possible and were confounded with the isolation versus context manipulation. However, the first three tasks were completed in the first session and the final task was completed in the second session. Thus, the reading tasks were completed on separate days. All tasks were separated by time and interleaved with filler tasks to minimise children's awareness of the repetition of items. The auditory lexical decision task was included to assess children's familiarity with the phonological forms (lexical phonology) and the definitions task was administered to tap item-specific lexical-semantic knowledge (lexical semantics).

Stimuli. Stimuli are included in the Appendix and comprised 20 regular words and 20 exception words, taken from longer lists in the Diagnostic Test of Word Reading Processes (DTWRP; Forum for Research in Literacy and Language, 2012). Regular words only included graphemes that were pronounced according to grapheme-phoneme correspondence (GPC) rules (Rastle & Coltheart, 1999), whereas exception words included one or more graphemes with pronunciations that deviated from these rules (e.g., the <s> in <sugar> has an atypical pronunciation). The stimuli included monosyllabic and multisyllabic words. Since stress patterns affect pronunciation in multisyllabic words, during DTWRP design an expert panel of psychologists, linguists and psycholinguists provided consensus that the regular multisyllabic words were pronounceable using usual grapheme-phoneme mappings. All words selected for the present study could be used as nouns. Regular and exception word lists were closely matched (all *p*s > .05) on length measured in phonemes, letters, or syllables, and on printed word frequency (where available from the Children's Printed Word Database; Masterson, Dixon, Stuart, & Lovejoy, 2003, otherwise from the CELEX Lexical Database; Baayen, Piepenbrock, & van Rijn, 1993). In addition, lists were matched (all *Fs* < 1) for bigram token frequency, bigram type frequency, trigram token frequency, trigram type frequency and number of orthographic neighbours (data from N-Watch; Davis, 2005). See Table 1 for a summary of the stimulus characteristics of the regular and exception words.

Table 1. Regular and exception word characteristics

Measure	Regular		Exception	
	M	*SD*	*M*	*SD*
Number of phonemes	5.90	1.74	5.10	1.55
Number of letters	7.05	1.76	6.30	1.95
Number of syllables	2.25	.85	2.00	.73
Printed word frequency[1]	140.21	206.70	174.81	348.83
Bigram frequency (token)[2]	1030.86	1339.56	1110.49	927.37
Bigram frequency (type)[2]	47.26	28.76	40.37	20.02
Trigram frequency (token)[2]	282.25	440.53	227.90	261.00
Trigram frequency (type)[2]	7.54	4.94	5.82	5.30
Orthographic N[2]	1.45	3.17	2.25	4.70

Notes. [1]Children's Printed Word Database (Masterson et al., 2003) and CELEX lexical database (Baayen et al., 1993); [2]N-Watch (Davis, 2005)

Reading tasks. In the first reading task, children read each word aloud in isolation. In the second, children read each word in a sentence context, with each word appearing at the end of a sentence stem ranging in length from four to nine words. In each trial of the contextualised reading task, a sentence stem was presented on the screen first. Following this, the target word was presented. Children were asked to read sentence stems and target words aloud. Sentence stems and target words were presented separately to minimise differences between the two reading tasks. In addition, the examiner corrected any errors made while reading sentence stems to maintain comprehension for the context. Errors made whilst reading target words were not corrected.

To develop sentence stems, regular and exception words were paired according to difficulty (using the difficulty order from the DTWRP; Forum for Research in Literacy and Language, 2012) so that sentence stems could be matched in pairs for overall printed word frequency (Masterson et al., 2003), and for length in words, letters and syllables (all *Fs* < 1). A series of cloze procedures was conducted with adults to develop contexts that were not overly constraining, such that participants could not readily guess the target from the sentence stem, and would therefore need to read it. For each cloze procedure adults were asked to complete each sentence stem (with target words missing). For the sentence stems used in this study, a maximum of 2/25 adults inserted the target in any one case, showing that children were unlikely to guess the target word from the sentence stem.

Within isolation and context reading tasks, trials were blocked by type (exception then regular). Stimuli were presented in random order within blocks using the E-Prime programme (Schneider, Eschman, & Zuccolotto, 2002a, 2002b). Words and sentences were presented in Arial 25-point font and the approximate viewing distance was 40cm. Words subtended an approximate mean visual angle of 4.37° to 10.82°, for four and ten letter words respectively. Accuracy was calculated for each child in each task (i.e., number of words read correctly). The maximum score was 20 for each list (regular, exception) within each task.

Auditory lexical decision. The auditory lexical decision task was administered to determine whether children were familiar with the phonological form of each word (lexical phonology). The 40 words were presented along with an equal number of nonwords from the ARC database (Rastle, Harrington, & Coltheart, 2002) that were matched to the words for number of letters and in most cases (80%) for initial phoneme. Items were recorded by a native speaker of English. Stimuli were presented one at a time through headphones and children were required to make a manual key-press response to indicate whether the item was a word or not. Children completed four practice trials at the beginning of the task to ensure they understood the task demands. Stimuli were presented in random order, and response accuracy (max = 20 for each word list) and latencies were recorded using E-Prime.

Definitions. Children were asked to describe what each word meant, yielding a measure of lexical-semantic knowledge. All 40 words were administered in a single random order. Items were blocked such that children responded to items from the exception word list first, and then items from the regular word list. The resulting definitions (N = 1080) were scored by two independent coders as 0 (no definition/incorrect definition), 1 (partial definition) or 2 (full definition). Criteria for scoring a 0, 1 or 2 for each word were agreed by the first author and coders beforehand. The coders then scored each definition without any consultation. There was a high degree of inter-rater reliability, $r(1080) = .96$. Nonetheless, the coders discussed each discrepant score in turn (with advice from the first author), reaching consensus in all cases. A total definitions score (max = 40 for each list) was calculated for each child.

Results

Mean normative scores were at or near the average range on standardised assessments of nonverbal reasoning, semantic knowledge and word-level reading (see Table 2 for a summary). High reliability estimates are reported for all tasks.

Table 2. Performance on standardised tasks

Measure (maximum raw score)	Reliability	Raw score M (SD)	Norm-referenced score M (SD)
Nonverbal reasoning (max raw score = 28)	.95[a]	9.26 (4.60)	50.30 (8.13)[1]
Vocabulary (max raw score = 56)	.87[a]	15.37 (4.1)	38.70 (7.15)[1]
Similarities (max raw score = 36)	.89[a]	13.26 (5.14)	51.26 (8.56)[1]
TOWRE PDE (max raw score = 63)	.90[b]	17.67 (13.36)	112.26 (14.46)[2]
TOWRE SWE (max raw score = 104)	.97[b]	40.96 (16.01)	114.67 (12.95)[2]

Notes. TOWRE = Test of Word Reading Efficiency; PDE = Phonemic Decoding Efficiency; SWE = Sight Word Efficiency; [a]Average split half reliability for 6-7 year olds according to the WASI manual; [b]Test/re-test reliability for 6-9 year olds according to the TOWRE manual; [1]T-score (*M* = 50, *SD* = 10); [2]Standard score (*M* = 100, *SD* = 15); maximum raw scores based on maximum number of items that could be administered to 6 – 8 year old children.

Table 3 summarises performance by participants and by items, and reliability estimates (Cronbach's α), for each experimental word task. Reliability estimates were acceptably high for most tasks, but were relatively low for auditory lexical decision.

Table 3. Performance on experimental tasks

Task	Condition	Reliability (Cronbach's α)	Performance by participants M (SD) [1]	Performance by items M (SD)[2]
Reading in isolation	Regular	.90	10.63 (4.84)	14.35 (8.06)
	Exception	.93	9.07 (5.71)	12.25 (7.03)
Reading in context	Regular	.90	11.48 (4.57)	15.50 (8.61)

	Exception	.91	11.22 (4.85)	15.15 (8.37)
Auditory lexical decision	Regular	.33	14.37 (1.90)	19.40 (7.94)
	Exception	.59	15.63 (2.34)	21.10 (6.49)
Definitions	Regular	.79	16.52 (5.15)	15.00 (10.45)
	Exception	.79	19.11 (5.60)	17.85 (8.21)

Notes. [1]Aggregated across participants (e.g., average number of regular words read correctly in isolation), for this, a maximum score for the definitions task within each condition = 40, for all other tasks a maximum score = 20; [2]aggregated across items (e.g., average number of participants reading regular words correctly in isolation), for all tasks the maximum number of participants = 27; to calculate Cronbach's alpha and performance by items for the definitions task a binary score was derived whereby an incorrect or no definition was coded 0 and partial or full definition coded 1.

We next present findings on: (i) correlation and regression analyses exploring general relationships between semantic knowledge and word-level reading (with scores calculated by participants in the more traditional way); and (ii) mixed-effects models that probe effects of regularity (regular vs. exception) and reading task (isolation vs. context), as well as item-specific relationships between semantic knowledge and word-level reading (taking into account random effects due to participants or items).

General relationships between semantic knowledge and word-level reading

Table 4 presents bivariate parametric correlations (by participants) between raw scores on standardised measures of semantic knowledge (vocabulary, similarities) and all reading tasks (TOWRE PDE, TOWRE SWE, regular and exception word reading in both isolation and context). Pertinent to our hypotheses, Table 4 indicates medium to large correlations between each measure of semantic knowledge and each measure of word-level reading. Contrary to our expectations, semantic variables were not more closely related to exception than regular word reading and, across word reading tasks, performance was less highly correlated with scores on the vocabulary than the similarities task. For nonword reading (measured by the TOWRE PDE), scores showed a higher correlation with vocabulary than similarities, though coefficients were similar.

Table 4. Parametric correlations (by participants) between standardised measures of semantic knowledge and word reading measures

Measure	1.	2.	3.	4.	5.	6.	7.	8.
1. Vocabulary	-							
2. Similarities	.41*	-						
3. TOWRE PDE	.52**	.48*	-					
4. TOWRE SWE	.42*	.61**	.86**	-				
5. Regular isolation	.44*	.67**	.81**	.85**	-			
6. Exception isolation	.44*	.61**	.86**	.91**	.88**	-		
7. Regular context	.47*	.64**	.82**	.91**	.90**	.91**	-	
8. Exception context	.52**	.60**	.78**	.85**	.86**	.89**	.96**	-

Notes. TOWRE = Test of Word Reading Efficiency; PDE = Phonemic Decoding Efficiency; SWE = Sight Word Efficiency; *$p < .05$; **$p < .01$

A series of regression analyses (see Table 5) was then conducted to probe whether semantic knowledge explains additional variance in word reading, after accounting for variance explained by phonological decoding ability (measured by TOWRE PDE score), which was entered at the first step. Separate analyses were conducted with performance on each word reading measure (number of words read correctly by each participant) as the outcome variable. Decoding was a significant independent predictor of each word reading measure, in each analysis. After accounting for the variance that decoding explained, similarities but not expressive vocabulary explained additional variance in each outcome variable. These models explained between 62% and 79% of the variance in word reading. Models with similarities explained more variance (67% – 79%), with similarities explaining about 5 – 10% of that variance. In summary, there was a clear relationship between semantic knowledge and word reading ability: this was more marked for the similarities task.

Table 5. Regression analyses predicting performance on word reading tasks (by participants) from nonword reading and semantic knowledge

	TOWRE SWE				Regular isolation				Exception isolation				Regular context				Exception context			
	B	S.E.	B	*p*	*B*	S.E.	β	*p*	*B*	S.E.	β	*p*	*B*	S.E.	β	*p*	*B*	S.E.	β	*p*
TOWRE PDE	1.05	.15	.88	<.001	.29	.05	.80	<.001	.37	.05	.85	<.001	.27	.05	.78	<.001	.25	.05	.69	<.001
Vocabulary	-.13	.47	-.03	.79	.04	.16	.03	.83	.01	.17	.01	.97	.08	.15	.07	.62	.19	.17	.16	.29
Model	R^2 = .74, p < .001				R^2 = .66, p < .001				R^2 = .73, p < .001				R^2 = .67, p < .001				R^2 = .62, p < .001			
TOWRE PDE	.89	.13	.74	<.001	.23	.04	.64	<.001	.31	.05	.73	<.001	.23	.04	.67	<.001	.23	.05	.63	<.001
Similarities	.77	.34	.25	.03	.35	.11	.37	<.001	.28	.12	.25	.03	.28	.11	.31	.01	.28	.13	.29	.04
Model	R^2 = .79, p < .001				R^2 = .76, p < .001				R^2 = .78, p < .001				R^2 = .74, p < .001				R^2 = .67, p < .001			

Notes. TOWRE = Test of Word Reading Efficiency; PDE = Phonemic Decoding Efficiency; SWE = Sight Word Efficiency

Mixed-effects models

In Generalized Linear Mixed-effects Models (GLMMs), we examined the factors that influenced the log odds of response accuracy, including fixed effects due to item regularity (regular vs. exception), experimental reading task (in isolation or in context), and word knowledge (auditory lexical decision or word definitions scores), as well as random effects due to variation in overall accuracy (random intercepts) or in the slopes of the fixed effects (random slopes) associated with differences between sampled participants or stimuli (Baayen, Davidson, & Bates, 2008). This approach allowed us to avoid the problems associated with analyzing dichotomous outcomes using linear models (discussed by e.g., Baayen, 2008; Dixon, 2008; Jaeger, 2008).

We analyzed 2160 observations - 27 children reading 20 regular and 20 exception words, once in each of the isolated and context conditions - using the glmer function in the lme4 package (Bates et al., 2014) in R (R Core Development Team, 2014). We tested the relative utility of including hypothesized fixed effects or potential random effects in our models by performing pair-wise Likelihood Ratio Test (LRT) comparisons (Barr, Levy, Scheepers, & Tily, 2013; Pinheiro & Bates, 2000) of simpler models with more complex models, where the former are nested within the latter. In the following, we outline the results of the model comparisons but report only estimates of fixed and random effects for the final model. Interested readers may examine the estimates associated with intermediate models in the Supplementary Materials, along with the data and the code used for all analyses.

First, we tested our hypotheses by progressing through a series of models with varying fixed effects but the same random effects, starting with a model of the log odds of response accuracy with no fixed effects and just the random effects of participants and items on intercepts (average accuracy) – an 'empty model'. Compared to the empty model, a model including terms corresponding to regularity, reading task, auditory lexical decision and definitions significantly improved model fit, LRT: $\chi^2 = 41.08$, 4 df, $p < .001$. In this main effects model, there were significant effects of reading task and definitions only (both *ps* < .001; see Supplementary Materials for full details). Our remaining hypotheses were addressed by adding interaction terms. Compared to the main effects model, a model also including the regularity x reading task interaction improved model fit, LRT: $\chi^2 = 5.30$, 1 df, $p = .021$. Adding regularity x auditory lexical decision and regularity x definitions terms did not further improve model fit, LRT: $\chi^2 = 0.47$, 2 df, $p = .789$. Thus, we adopted a final model that included the main effects and regularity x reading task terms.

Following Baayen (2008; see also Pinheiro & Bates, 2000), we examined whether both random intercepts terms were required by performing pairwise LRT comparisons of models with the same fixed effects as the final model but varying random effects as follows: (i) a model with both random effects of participants and items on intercepts, as in the models detailed in the foregoing; compared to (ii) a model with just the random effect of participants on intercepts; and compared to (iii) a model with just the random effect of items on intercepts. We found that both random intercepts terms were warranted by improved model fit to data (inclusion of a random effect of participants on intercepts, LRT: $\chi^2 = 754.85$, 1df, $p < .001$; inclusion of a random effect of items on intercepts, LRT: $\chi^2 = 585.05$, 1df, $p < .001$). In models (ii) and (iii) the pattern of significant effects remained largely the same, with significant effects of reading task and definitions and a regularity x reading task term that was near significant (model ii: p = .058; model iii: p = .094). However, in each model, the auditory lexical decision effect was additionally significant (both *ps* < .001, see supplementary materials for further details). Thus, when variation relating to either participants or items alone was taken into account, both auditory lexical decision and definitions showed a significant relationship with word reading. However, after simultaneously accounting for variation relating to both, only the definitions effect remained.

Following Barr et al.'s (2013) recommendations, we examined the importance of random slopes (random differences between participants or between items in the slopes) of the fixed effects due to reading task, word knowledge or the regularity x reading task interaction. We did this by testing whether model fit was improved by the inclusion of terms corresponding to random effects of participants or items on the slopes of the fixed effects. We found that a model including terms corresponding to random effects of participant differences on the slopes of word regularity and reading task effects, and corresponding to random effects of item differences on the slopes of both word knowledge measures (definitions and auditory lexical decision), significantly fit the data better than a model including the same fixed effects and just random intercepts, LRT: χ^2 = 34.46, 10df, p < 0.01. Thus, including the observed variability between participants in the slopes of both regularity and reading task effects, and between responses to different items in the slope of the word knowledge effect, improved model fit.

Table 6 summarises the final model, with fixed effects due to regularity, reading task, the regularity x reading task interaction, and word knowledge (scores on auditory lexical decision and definitions tests), as well as random effects of participants and items on intercepts, and on the slopes of the fixed effects.

Table 6. Summary table of the Generalised Linear Mixed-effects model of word reading

Fixed effects	Estimated coefficient	*SE*	*z*	*p*
(Intercept)	.34	.90	.37	.71
Item regularity (regular vs. exception)	.09	.61	.15	.88
Reading task (isolated vs. context)	-1.23	.25	-4.98	<.001
Word knowledge (auditory lexical decision)	.13	.32	.40	.69
Word knowledge (definitions)	.33	.14	2.31	.02
Item regularity x reading task interaction	.79	.29	2.71	<.01
Random effects	**Variance**	***SD***	**Correlation**	
Due to items				
Intercepts	16.02	4.00		
Word knowledge (auditory lexical decision)	1.65	1.28	-.99	
Word knowledge (definitions)	.19	.43	-.94	.88
Due to participants				
Intercepts	7.03	2.65		
Item regularity (regular vs. exception)	.31	.56	-.87	
Reading task (isolated vs. context)	.39	.63	.26	.24

Note. Number of observations: 2160; 40 items; 27 participants. Correlations are the estimated correlations between Best Linear Unbiased Predictors (the random effects).

The estimated coefficients for the final model show that reading accuracy was higher for the context (vs. isolation) task, and for words that had been defined more accurately. Further, the model revealed a regularity x reading task interaction. Inspection of Table 3 indicates that a regularity effect was more evident when words were read in isolation rather than in context, and that the influence of context was greater for exception than regular words. Contrary to our hypotheses, we found that 1) semantic knowledge showed equivalent relationships with regular and exception word reading, and 2) auditory lexical decision performance was not associated with word reading.

Discussion

The results support our primary hypothesis that variation in semantic knowledge is associated with variation in word reading performance. Indeed, we have provided robust evidence for this by observing this association, for the first time, across both general and item-specific analyses.

We have extended previous findings on reading words in isolation by assessing word reading in context, which is more akin to how children encounter words naturally. We observed an interaction between context and word type such that sentence context particularly facilitated reading of exception words, in line with previous studies (Nation & Snowling, 1998). It is worth noting that the reading in isolation task was always administered before the reading in context task and thus any contextual benefit must be interpreted with caution as some improvement might be attributable to practice effects. However, tasks were separated by approximately one week. Further, the order of the tasks does not invalidate our finding of an interaction between context and word type, nor does it impact on our key findings that semantic knowledge (as measured by the similarities task) predicted reading in context as well as reading in isolation in regression analyses, and that lexical-semantic knowledge and contextual effects were independently predictive of word reading in our mixed-effects analyses. Theories of word reading focus almost exclusively on reading in isolation. Nonetheless, our findings are consistent with developmental theories that highlight the importance of contextual support for word reading (Share, 1995) and with the triangle model's (yet to be implemented) assumption that semantics and context exert separable but interacting effects on reading aloud (Seidenberg & McClelland, 1989; Bishop & Snowling, 2004). We hope that the present study, along with other empirical studies of word reading in context (Martin-Chang & Levesque, 2013; Nation & Snowling, 1998) will pave the way for research that aims to probe the mechanisms that underpin word reading as it occurs naturally. An important first step will be to specify how context supports word reading, why this might be more beneficial for exception than regular word reading, and why this effect was separable from that of item-specific semantic knowledge in our analyses.

In the present study, context was provided at the sentence level and may have conveyed useful semantic information, along with other cues (e.g., grammar). Thus, one plausible interpretation of our findings would be that semantic information from the context supported word reading, and this was more effective for exception than regular words. However, this interpretation is premature; our data do not address whether this effect was driven by semantic information or other cues provided by context. We found that semantic knowledge showed equivalent relationships with regular and exception word reading, a finding that we replicated across by-participants regression analyses and mixed-effects models. To the extent that our measures of semantic knowledge map onto the way that semantic representations are activated in the triangle model, this finding contrasts with the triangle model, where semantic knowledge is seen as more important for exception word reading than regular word reading (e.g., Harm & Seidenberg, 2004; Strain, Patterson, & Seidenberg, 1995; though see Woollams et al., 2007 for effects of semantics on regular word reading within a triangle model framework). Notably, it is also at odds with pertinent developmental findings that semantic knowledge shows a closer relationship with exception word reading than regular word reading in English-speaking children (Nation & Cocksey, 2009; Ricketts et al., 2007), whilst according with emergent findings from English-speaking children indicating relationships between semantic variables and both regular and exception word reading (Duff & Hulme, 2012; Mitchell & Brady, 2013, see also findings from Spanish-speaking adults, reported by Davies, Barbón, & Cuetos, 2013, and English-speaking adults, reported by Strain & Herdman, 1999). It remains to be seen whether this finding is predicted by the DRC model (Coltheart et al., 2001) as current instantiations have not yet simulated the role of semantics in word reading development (see Taylor, Rastle, & Davis, 2013).

There are a number of possible explanations for discrepancies between our observations and previous findings. Marked ceiling effects on regular word reading could explain weaker relationships between semantic knowledge and regular word reading in previous studies (Nation & Cocksey, 2009; Ricketts et al., 2007). Another possibility concerns the age and reading ability of participants. Semantic knowledge may contribute more indiscriminately to word reading in the early stages of reading development when children have limited knowledge of orthography-to-phonology mappings (as in our study; for a similar argument, see Duff & Hulme, 2012). With reading experience, the role of semantics in regular word reading may reduce, such that a closer relationship between semantic knowledge and exception word reading emerges. Additionally, the impact of semantic knowledge on word reading may be influenced by item-level characteristics such as length, frequency, familiarity and meaning (Mitchell & Brady, 2013). Indeed, our set of regular words were harder to define than our set of exception words. This could go some way to explaining the finding that semantic knowledge contributes to both regular and exception words. Future research should aim to explore the conditions under which semantic knowledge impacts on regular word reading, adopting developmental designs and varying stimulus characteristics.

In correlation analyses (by participants) all standardised measures of semantic knowledge and word-level reading were inter-correlated. However, knowledge of semantic relationships (similarities) was consistently more highly correlated with word reading than oral vocabulary knowledge. After controlling for decoding skill, regression analyses showed that similarities but not expressive vocabulary predicted word reading. One possible explanation for this finding is that scores on the similarities measure were more varied than scores on the vocabulary measure such that the similarities measure may have captured more fully the variability in semantic knowledge in our sample. The hypothesis that performance on the similarities measure was systematically more varied than performance on the oral vocabulary measure could be explored in future research. Previous studies that have investigated general relationships between more than one semantic measure and word reading have shown that the semantic predictors of word reading (after controlling for decoding) vary according to the age of the participants and the outcome measures used in analyses. In Ouellette and Beers (2010) both depth and breadth of vocabulary knowledge was measured. In younger participants (5-7 years) depth but not breadth predicted irregular word reading whereas in older participants (11-12 years) the opposite pattern was observed. Nation and Snowling (2004) employed a measure of vocabulary and a 'semantic composite' (semantic fluency and synonym judgement). Both measures predicted word reading concurrently but only oral vocabulary was a longitudinal predictor of exception word reading.

The finding that oral vocabulary did not predict word reading in our regression analyses also contrasts with the item-specific effects detected in our mixed-effects models. This seems surprising given that the definitions tasks was designed to parallel the standardised expressive vocabulary measure that we used by asking children to define words and adopting a three-point scoring approach. Plausibly, this discrepancy could be explained by differences in the variables included in the models. We controlled for decoding ability in our by-participants analyses so that we could examine the relationship between semantic knowledge and word reading after accounting for the substantial variance in word reading explained by decoding skill (this is a standard approach, see for example Ouellette & Beers, 2010; Ricketts et al., 2007). However, we did not include decoding ability in our mixed-effects analyses because the models were specified as confirmatory analyses (following e.g., Barr et al., 2013) of the effects of the following experimental factors: reading task, regularity, and word knowledge type. Nevertheless, the addition of decoding ability to the final model did not change the pattern of results (see Supplementary Materials for details). Different findings across our analytical approaches could instead reflect the way that our mixed-effects models capture a specific relationship between knowledge of an item and reading that same item, whereas the by-participants regressions explore a more general relationship between a measure of a child's lexical-semantic knowledge, which could act as a proxy for their item-specific semantic knowledge or their ability to use context, and their ability to read a separate set of words. Arguably, this general relationship could be weaker. Taken together with the mixed findings discussed in our preceding paragraph it is clear that while the relationship between semantic knowledge and word reading is robust, the precise pattern of findings observed varies across analyses and data sets. Notably though, our observations show that semantic knowledge is predictive of word reading ability.

Mixed-effects models demonstrated that correctly defining a word was a significant predictor of accurately reading that word whereas accepting it as a word in our lexical decision task was not. This result was unexpected given that in Nation and Cocksey (2009), performance on definitions and auditory lexical decision tasks showed equivalent (significant) correlations with word reading and that auditory lexical decision was the stonger predictor in by-items regression analyses (for similar findings, see Duff & Hulme, 2012, Experiment 2; McKague et al., 2001). Thus, we did not replicate Nation and Cocksey's finding that auditory lexical decision predicts word reading, nor did we provide support for their proposal that lexical phonology is enough to support word reading (i.e. lexical-semantic knowledge provides no additional benefit). Instead, our findings indicate that it is lexical-semantic rather than lexical-phonological knowledge that supports word reading. Other investigations of the relative importance of lexical phonology and semantics for word reading have indicated that semantic knowledge is a better predictor than phonological knowledge of reading success (Duff & Hulme, 2012, Experiment 1; McKay et al., 2008; Taylor et al., 2011), resonating with our findings.

One plausible explanation for the discrepancy between our study and that of Nation and Cocksey (2009) relates to the different analytic approaches adopted in the studies. In Nation and Cocksey, correlations and regressions were conducted across items (an F2 by-items analysis), thus taking into account random error variance due to the items. In contrast, our final mixed-effects model incorporated random error variance due to both participants and items. Thus, accounting for both sources of error variance could have 'washed out' the effect of auditory lexical decision. Indeed, when our model accounted for either error variance due to participants (akin to F1 by-participants analyses) or error variance due to items (akin to F2 by-items analyses), we replicated the Nation and Cocksey finding: both definitions and auditory lexical decision performance predicted word reading. Analyses reported by Baayen et al. (2008) indicate that fixed effects are better estimated in repeated measures studies when both random participant and item effects are taken into account (see also Barr et al., 2013). Essentially these models specify, rather than assume, the random variation in the data that is due to participants (in this case variation in children's reading accuracy) and items (in this case variation in performance in response to individual words). It is possible that our findings would be replicated in Nation and Cocksey's data if mixed-effects models were applied, supporting a conclusion that lexical semantics but not lexical phonology impacts on word reading.

Caution is warranted in interpreting our auditory lexical decision results. Reliability for this task was low and post-hoc consideration of its stimuli has highlighted its limitations. Following Nation and Cocksey (2009) we selected nonwords that matched our words in terms of letter length and initial letter (or phoneme). However, we should have explicitly matched words and nonwords for number of syllables and phonemes. We checked this retrospectively, discovering that our words had approximately one more phoneme (M = 5.50, SD = 1.68 vs. M = 4.45, SD = 1.11) and one more syllable (M = 2.13, SD = .79 vs. M = 1.03, SD = .16). It is possible that this made the nonwords superficially distinctive from the words, making the task easier and reducing the extent to which lexical knowledge was used to make decisions (they could instead have been made on the basis of shallower processing). By participants there is no indication of ceiling effects and performance showed good variability. By items, performance again showed good variability but scores were closer to ceiling (this is also the case in Nation & Cocksey, 2009), providing some evidence that discriminating between particular words and nonwords was fairly easy. Ceiling effects by items may also explain poor reliability (Cronbach's alpha) on the auditory lexical decision task. Our choice of nonword distracters may therefore have restricted relationships between auditory lexical decision performance and reading because auditory lexical decision performance did not consistently reflect lexical knowledge or because scores on this task showed poor reliability (for further discussion of the impact of poor reliability on correlational analyses, see Vul, Harris, Winkielman, & Pashler, 2009).

The nature of the nonwords used in auditory lexical decision tasks has important implications for how performance on this task should be interpreted (e.g., Ernestus & Cutler, 2015). As mentioned above, superficial differences between our word and nonword stimuli may have reduced the use of lexical knowledge in making decisions. Equally though, in tasks where nonwords are very word-like, lexical decisions are commonly assumed to reflect greater reliance on semantic processing (Binder et al., 2003). An important goal for future research will be to investigate the relative contributions of lexical phonology and semantic knowledge to word reading using more carefully controlled auditory lexical decision stimuli and/or other tasks designed to tap lexical phonology.

In sum, our findings provide robust and novel support for the idea that semantic knowledge and sentence context independently support word reading (cf. Bishop & Snowling, 2004). In addition, they add to emergent evidence that lexical or semantic knowledge supports reading of regular as well as exception words (Davies, Barbón, & Cuetos, 2013). If semantic knowledge is causally related to word reading success then training knowledge of word meanings should benefit word reading. Findings from such training studies have so far been inconclusive, with some suggesting that training lexical-level phonological knowledge is sufficient to support word reading (Duff & Hulme, 2012, Experiment 2; McKague et al., 2001) and others indicating that semantic knowledge exerts an effect beyond phonology (McKay et al., 2008; Taylor et al., 2001). Future empirical and theoretical studies that adopt psychologically plausible approaches to learning and development should aim to advance our understanding of how the relationship between lexical knowledge and word reading changes with age and development, and whether semantic knowledge is causally related to word reading.

References

Archer, N., & Bryant, P. (2001). Investigating the role of context in learning to read: A direct test of Goodman's model. *British Journal of Psychology, 92*, 579-591. doi: 10.1348/000712601162356

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York: Cambridge University Press.

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*(4), 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, University of Pennsylvania.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278. doi: 10.1016/j.jml.2012.11.001

Binder, J., McKiernan, K., Parsons, M., Westbury, C., Possing, E., Kaufman, J., & Buchanan, L. (2003). Neural correlates of lexical access during visual word recognition. *Journal of Cognitive Neuroscience, 15*(3), 372-393. doi:10.1162/089892903321593108

Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130*(6), 858-886. doi: 10.1037/0033-2909.130.6.858

Bowey, J. A., & Rutherford, J. (2007). Imbalanced word-reading profiles in eighth-graders. *Journal of Experimental Child Psychology, 96*(3), 169-196. doi: 10.1016/j.jecp.2006.11.001

Brady, S., & Shankweiler, D. (1991). *Phonological processes in literacy: A tribute to Isabelle Y. Liberman*. Hillsdale, NJ: Erlbaum.

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading comprehension difficulties: A randomised controlled trial. *Psychological Science, 21*, 1106-1116. doi: 10.1177/0956797610375449

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. Psychological Review, 108 , 204-256. doi: 10.1037/0033-295X.108.1.204

Davies, R., Barbón, A., & Cuetos, F. (2013). Lexical and semantic age-of-acquisition effects on word naming in Spanish. *Memory & Cognition, 41*(2), 297-311. doi: 10.3758/s13421-012-0263-8

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods, 37*, 65-70. doi: 10.3758/BF03206399

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language, 59*(4), 447-456. doi: 10.1016/j.jml.2007.11.004

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*(6), 504-525. doi: 10.1080/10888438.2011.598199

Ernestus, M., & Cutler, A. (2015). BALDEY: A database of auditory lexical decisions. *The Quarterly Journal of Experimental Psychology, 68*(8), 1469-1488. doi:10.1080/17470218.2014.984730

Forum for Research in Literacy and Language (2012). *Diagnostic Test of Word Reading Processes (DTWRP)*. London: GL Assessment.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hove, East Sussex: Psychology Press Ltd.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6-10. doi: 10.1177/074193258600700104

Graham, K. S., Hodges, J. R., & Patterson, K. (1994). The relationship between comprehension and oral reading in progressive fluent aphasia. *Neuropsychologia, 32*, 299-316. doi: 10.1016/0028-3932(94)90133-3

Harm, M., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*(3), 662-720. doi: 10.1037/0033-295X.111.3.662

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434-446. doi: 10.1016/j.jml.2007.11.007

Martin-Chang, S., & Levesque, K. (2013). Taken out of context: differential processing in contextual and isolated word reading. *Journal of Research in Reading, 36*(3), 330-349. doi:10.1111/j.1467-9817.2011.01506.x

Masterson, J., Dixon, M., Stuart, M., & Lovejoy, S. (2003). The Children's Printed Word Database. from http://www.essex.ac.uk/psychology/cpwd/

McKague, M., Pratt, C., & Johnston, M. B. (2001). The effect of oral vocabulary on reading visually novel words: A comparison of the dual-route-cascaded and triangle frameworks. *Cognition, 80*, 239-270. doi: 10.1016/S0010-0277(00)00150-5

McKay, A., Davis, C., Savage, G., & Castles, A. (2008). Semantic involvement in reading aloud: Evidence from a nonword training study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1495-1517.

Mitchell, A. M., & Brady, S. A. (2013). The effect of vocabulary knowledge on novel word identification. *Annals of Dyslexia, 63*(3-4), 201-216. doi: 10.1007/s11881-013-0080-1

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*(3), 296-308. doi: 10.1016/j.jecp.2009.03.004

Nation, K., & Snowling, M. J. (1998). Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child Development, 69*(4), 996-1011. doi: 10.1111/j.1467-8624.1998.tb06157.x

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*, 342-356. doi: 10.1111/j.1467-9817.2004.00238.x

Ouellette, G. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554-566. doi: 10.1037/0022-0663.98.3.554

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, 23*(2), 189-208. doi: 10.1007/s11145-008-9159-1

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*(2), 239-251. doi: 10.1016/j.jecp.2009.05.001

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Plaut, D. C., McClelland, J. L., Seidenberg, M., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56-115.

Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 25*(2), 482-503. doi:10.1037/0096-1523.25.2.482

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *The Quarterly Journal of Experimental Psychology Section A, 55*(4), 1339-1362. doi: 10.1080/02724980244000099

Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading, 11*(3), 235-257. doi: 10.1080/10888430701344306

Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools Inc.

Schwartz, M.F., Saffran, E.M. & Marin, O.S.M. (1980). Fractioning the reading process in dementia: Evidence for word-specific print-to-sound associations. In M. Coltheart, K. E. Patterson, & J. C. Marshall (Eds.). *Deep Dyslexia*. London: Routledge & Kegan Paul.

Seidenberg, M., & McClelland, J. L. (1989). A distributed, developmental model of word recognition. *Psychological Review, 96,* 523-568. doi: 10.1037/0033-295X.96.4.523

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151-218. doi: 10.1016/0010-0277(94)00645-2

Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology, 53*(4), 347 – 359. doi: 10.1037/h0087322

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(5), 1140-1154. doi:10.1037/0278-7393.21.5.1140

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*(4), 322-328. doi:10.1177/0963721415574980

Taylor, J., Plunkett, K., & Nation, K. (2011). The influence of consistency, frequency, and semantics on learning to read: An artificial orthography paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 37*(1), 60-76. doi: 10.1037/a0020126

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin, 139*(4), 766-791. doi: 10.1037/a0030266

Torgesen, J., Wagner, R., & Rashotte, C. (1999). *Test of Word Reading Efficiency*. Austin, TX: Pro-Ed.

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*(2), 122-140. doi: 10.1080/10888438.2010.542527

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science, 4*(3), 274-290. doi:10.1111/j.1745-6924.2009.01125.x

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing: An Interdisciplinary Journal, 16*, 5-20. doi: 10.1023/A:1021789804977

Wang, H.-C., Nickels, L., Nation, K., & Castles, A. (2013). Predictors of orthographic learning of regular and irregular words. *Scientific Studies of Reading, 17*(5), 369-384. doi:10.1080/10888438.2012.749879

Woollams, A. M., Ralph, M. A. L., Plaut, D. C., & Patterson, K. (2007). SD-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review, 114*(2), 316-339. doi:10.1037/0033-295X.114.2.316

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. London: The Psychological Corporation.

Appendix. Experimental word stimuli

Regular words	Exception words
dragon	ball
well	monkey
mouse	half
elephant	ghost
street	many
corner	sugar
kettle	want
noise	giant
ostrich	island
chimpanzee	station
picnic	soup
goblin	cousin
banister	stomach
statue	vehicle
marzipan	restaurant
turmoil	parachute
sacrifice	reservoir
wilderness	mosquito
auditorium	sovereign
anecdote	horizon

Highlights

- Relationships between semantic knowledge and word reading were explored.
- Data were analysed using regression approaches and mixed-effects models.
- Semantic knowledge predicted regular and exception word reading in six year olds.
- Separately, there was an additional positive effect of reading words in context.
- The findings support a role for semantic knowledge and context in word reading.