THE ROLE OF SEMANTICS IN ORTHOGRAPHIC AND PHONOLOGICAL LEARNING

A dissertation submitted to Kent State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

by

Stephen M. Brusnighan

May, 2015

© Copyright

All rights reserved

Except for previously published materials
Dissertation written by

Stephen M. Brusnighan

B.S., High Point University, 2008
M.A., Kent State University, 2010
Ph.D., Kent State University, 2015

Approved by

________________________________, Advisor
Jocelyn Folk, Ph.D., Psychological Sciences

________________________________, Member
William Merriman, Ph.D., Psychological Sciences

________________________________, Member
Katherine Rawson, Ph.D., Psychological Sciences

________________________________, Member
Mary Beth Spitznagel, Ph.D., Psychological Sciences

________________________________, Outside Member
Anna Schmidt, Ph.D., CCC/SLP, Speech Pathology & Audiology

________________________________, Graduate Representative/Moderator
Sarah Rilling, Ph.D., English

Accepted by

________________________________, Chair, Department of Psychological Sciences
Maria Zaragoza, Ph.D.

________________________________, Dean, College of Arts and Sciences
James Blank, Ph.D.
# TABLE OF CONTENTS

TABLE OF CONTENTS .................................................................................................................. iii

LIST OF TABLES .......................................................................................................................... vi

ACKNOWLEDGMENTS ................................................................................................................. vii

INTRODUCTION .......................................................................................................................... 1

  Lexical Quality in Orthographic and Phonological Learning .................................................... 2

  Effects of Contextual Constraint and Reading Exposure in Reading Novel Words ...................... 8

  Post-reading Tests of Semantic Learning from Incidental Exposure during Silent Reading ........ 9

  The Current Study ..................................................................................................................... 12

EXPERIMENT 1 ............................................................................................................................. 15

  Method ...................................................................................................................................... 16
    Participants ................................................................................................................................. 16
    Materials .................................................................................................................................... 16
      Target words .............................................................................................................................. 16
      Sentence contexts .................................................................................................................... 17
      Spelling recognition test .......................................................................................................... 19
      Meaning recognition vocabulary test ..................................................................................... 20
    Procedure ................................................................................................................................... 21
    Apparatus .................................................................................................................................. 23
  Results 1a: Eyetracking .............................................................................................................. 23
    Effects of contextual constraint in reading novel words .......................................................... 24
  Results 1b: Post-reading Tests of Word Learning ....................................................................... 28
    Meaning recognition vocabulary test results .......................................................................... 29
    Spelling recognition test results .............................................................................................. 31
    Response comparisons .............................................................................................................. 32
    Naming time task results .......................................................................................................... 33
Discussion .................................................................................................................. 34

EXPERIMENT 2 ........................................................................................................... 39

Method .......................................................................................................................... 41
  Participants ............................................................................................................... 41
  Materials .................................................................................................................. 42
  Procedure ................................................................................................................ 43
  Apparatus ................................................................................................................ 44
  Results 2a: Self-paced Reading .............................................................................. 44
  Results 2b: Post-reading Tests of Word Learning .................................................. 45
    Meaning recognition vocabulary test results ...................................................... 45
    Spelling recognition test results ....................................................................... 47
    Response comparisons ....................................................................................... 49
    Naming time task results ................................................................................... 50
  Discussion ............................................................................................................... 53

GENERAL DISCUSSION ............................................................................................. 56

  Lexical Quality in Orthographic and Phonological Learning ............................... 56

  Effects of Contextual Constraint and Reading Exposure in Reading Novel Words ............................................................................................................. 58

  Post-reading Tests of Semantic Learning from Incidental Exposure during Silent Reading .............................................................................................................. 59

APPLICATIONS AND FUTURE DIRECTIONS .............................................................. 61

REFERENCES .......................................................................................................... 64

APPENDIX A .............................................................................................................. 74

APPENDIX B .............................................................................................................. 82

APPENDIX C .............................................................................................................. 83
LIST OF TABLES

TABLES

1 Mixed-effects Models for Target Word Processing in Experiment 1.
2 Means for Reading Times on the Target Word by Condition in Experiment 1.
3 Total Reading Times on Target Words for Repeated Items by Trial Order in Experiment 1.
4 Mixed-effects Models for the Post-reading Tests of Learning in Experiment 1.
5 Vocabulary Test Accuracy by Number of Exposures in Experiment 1.
6 Means for Spelling Recognition Accuracy in Experiment 1.
7 Means for Spelling Recognition Test Responses by Exposure Condition in Experiment 1.
8 Mean Naming Times and Pronunciation Accuracies in Experiment 1.
9 Means for Token Bigram Frequencies of Spelling Recognition Test Responses in Experiment 1.
10 Mixed-effects Models for Reading on the Target Word in Experiment 2.
11 Mean Target Word Reading Times by Trial Order in Experiment 2.
12 Vocabulary Test Accuracy by Context Condition in Experiment 2.
13 Spelling Accuracy and Target Word Reading Time as a function of Vocabulary Test Accuracy in Experiment 2.
Vocabulary Test Accuracy as a function of Accuracy on the Spelling Test in Experiment 2.

Means for Spelling Recognition Test Responses in Experiment 2.

Naming Times and Pronunciation Accuracy by Context Condition in Experiment 2.

Naming Times for Novel Words by Participants in Experiment 2 versus in a Control Group.

Means for Token Bigram Frequencies of Spelling Recognition Test Responses in Experiment 2.
ACKNOWLEDGMENTS

This project would not have been possible without the support, guidance, and helpful suggestions of several people. I would particularly like to thank Jocelyn Folk for advising me on this project. I would also like to thank the members of my committee: William Merriman, Katherine Rawson, Mary Beth Spitznagel, and Anna Schmidt.
INTRODUCTION

Given the importance of vocabulary knowledge in reading comprehension skill (National Institute of Child Health and Human Development, 2000), it is crucial to understand the processes involved in incidental vocabulary learning from reading. After all, once children have learned to read, most new vocabulary words are learned through incidental exposure during normal reading (Carlisle, 2007; Drum & Konopak, 1987; Durkin, 1978; Jenkins & Dixon, 1983; Landauer & Dumais, 1997; Long & Shaw, 2000; Nagy, Herman, & Anderson, 1985; Shu, Anderson, & Zhang, 1995; Sternberg, 1987). Chaffin, Morris, and Seely (2001) estimated that people ages 2-20 learn between 6 and 25 words per day. As such, a common experience during normal reading is encountering a new word several times in a passage. When a new word appears multiple times in a text, it appears in a number of different sentence contexts from which readers may be able to infer a meaning for the novel word. The sentence context cues used in each of these encounters with new words span the continuum from misleading to highly informative as to what the words mean. Studies investigating multiple exposures to new vocabulary words and the nature of contextual constraint in vocabulary learning have mostly focused on semantic learning (Bolger, Balass, Landen, & Perfetti, 2008; Durso & Shore, 1991; Frishkoff, Collins-Thompson, Perfetti, & Callan, 2008; Frishkoff, Perfetti, & Collins-Thompson, 2010; Frishkoff, Perfetti, & Westbury, 2009). That is, the goal of these studies has been to investigate changes in knowledge states about the meaning of words over several exposures in a variety of contexts. With the goal of these investigations
being *semantic learning* the other aspects of word learning are often neglected such as learning a word’s orthography (representation of a word’s spelling) and phonology (representation of a word’s sound). The goal of the current study was to investigate, over multiple exposures to novel words in a variety of contexts, the nature of *orthographic* and *phonological learning* and the role of semantics in these types of learning by skilled, adult readers. Specifically, this study provides an account of incidental word learning from reading that quantifies (1) the relationship between semantic learning and the learning of word form, (2) the effects of contextual constraint on reading times for novel monomorphemic words during reading in English, and (3) the amount of semantic, orthographic, and phonological learning possible from a single reading encounter vs. from several encounters with a novel word in the same reading session.

**Lexical Quality in Orthographic and Phonological Learning**

The Lexical Quality Hypothesis (LQH, Perfetti, 2007; Perfetti & Hart, 2001; 2002) posits that lexical representations vary in quality and that these variations in lexical quality (LQ) affect skilled and less-skilled readers differently such that skilled readers tend to have higher quality lexical representations. That is, high LQ representations of words are said to have fully specified forms (phonology and orthography) and meanings (semantics). In the same way, low LQ representations of words are said to contain only partial information such as orthography and/or phonology without semantics. Applied to word learning from reading, this hypothesis would imply that when all three types of information are available in a word learning episode from which to form representations (i.e., phonology, orthography, and semantics are encoded together) these representations
have an opportunity to be of high LQ and would only strengthen with additional exposures (Reichle & Perfetti, 2003). In contrast, if only partial information is available, such as form without meaning, representations would be of lower LQ. Balass, Nelson, and Perfetti (2010) provided Event Related Potential (ERP) evidence for this account in adult skilled readers. They found that learning of a word’s phonology and learning of a word’s orthography both benefit from learning of that word’s meaning. This ERP evidence suggests that reading words in a constraining linguistic context would lead to better orthographic learning and better phonological learning than reading words in neutral text. However, they used an explicit learning task. To test whether this applies to an incidental learning situation, evidence is needed from surprise tests of orthographic and phonological learning following a reading session in which readers encounter novel words both in constraining and neutral sentence contexts.

The LQH also specifies that for learning words from reading the lexical quality of word representations builds incrementally with each new encounter such that after several exposures, word knowledge would presumably have a better chance of being robust than after a single exposure. This notion has been tested using direct instruction methods (Frishkoff et al., 2008) such that semantic knowledge improves over several exposures, but this has not been tested using incidental vocabulary acquisition methods during silent reading. Regarding word form, Perfetti (2007) posits that representations of word form build incrementally as well, such that a representation of a word’s spelling would build stability over exposures, becoming more and more specified until it overlaps with the correct spelling. Again, to test this directly, evidence is needed from surprise
tests of orthographic, phonological, and semantic learning following a reading session in which novel words are encountered incidentally. It may be the case that, in an incidental task for which the goal of reading is comprehension, rather than word learning, readers do not demonstrate any more semantic knowledge for new words encountered several times in context than words encountered once in context. However, if representations of form and meaning also build incrementally during an incidental reading task, performance on post-reading tests of learning for form and meaning should be greater for words seen over several exposures rather than in one exposure.

Previous studies have provided tests of orthographic learning, but the literature regarding this issue has primarily involved children as participants (Bowey & Miller, 2007; Castles & Nation, 2006; Cunningham, 2006; Cunningham, Perry, Stanovich & Share, 2002; Ehri, 1993; 1995; 2005; Ehri & Wilce, 1980; 1987; Lucas & Norbury, 2014; Nation, Angell, & Castles, 2007; Ricketts, Bishop & Nation, 2008; Ricketts, Bishop, Pimperton, & Nation, 2011; Share, 1995; 1999; 2004; Treiman, 1993; Wimmer & Goswami, 1994; Wimmer, Mayringer, & Landerl, 2000; Ziegler & Goswami, 2005).

Webb (2008) found that for Japanese adults (university students) learning English as a foreign language, context quality has a greater impact on semantic knowledge than number of reading exposures, while number of reading exposures has a greater impact on learning of word form than context quality. In a study of adults who were native speakers of English, Bolger et al. (2008) investigated semantic learning and orthographic learning of unknown words encountered during reading in either 1 or 4 varied or repeated contexts. Evidence from two experiments was mixed with regard to differences in
orthographic learning based on the number and variety in contexts. The data were also mixed as to the influence of semantic learning on orthographic learning. Specifically, evidence that orthographic knowledge over 4 exposures was better than after 1 exposure was present in Experiment 1 but absent in Experiment 2. Evidence that orthographic knowledge was better after new words were seen in informative contexts than in neutral sentences was absent in Experiment 1 but present in Experiment 2.

Further research testing incidental orthographic learning in adults from reading in English (Brusnighan, Morris, Folk, & Lowell, 2014) has shown that previously used methods may demonstrate little about memory for orthography as they are confounded with episodic memory for the task material. In Brusnighan et al.’s study, a surprise cued recall test was given after a reading session in which participants’ eye movements were monitored as they read novel and known words during reading. For example, readers saw the sentence pair:

The musician clanged the chenth three times.

The instrument added excitement to the concert.

Then, on the cued recall test, in an attempt to have participants produce a spelling for each novel word from memory, they were given a cue:

“Write as much as you remember about ‘the musician at the concert’.”

In the cued recall data from the experiment, approximately 63% of novel word cues were left blank and 8% of cues received a response such as “I don’t remember.” When readers saw a known word such as gong in place of chenth in the reading portion of the study, for example, the majority of cues were still left blank (51%). While 32% of
the novel words were spelled correctly on the trials that were not left blank, it is evident that the cued recall method is confounded with memory for the task material. That is, it is not necessarily the case that participants could not remember the correct spelling for a word, but based on the memory cue, they could often not remember the content of any of the rest of the sentence. Given that cued recall tests for novel word spellings have been confounded with episodic memory for task materials (and in turn have very low response rates), it is entirely possible that one reading exposure is not sufficient for readers to demonstrate knowledge of a new word’s form. As such, it might take several exposures to new vocabulary words in a variety of contexts in order to demonstrate learning (with appropriate tests of learning that are not confounded with memory for the task material).

Previous research suggests that post-reading tests of spelling recognition demonstrate more about readers’ orthographic learning than cued recall methods. Although evidence is from intentional rather than incidental word learning tasks, performance on such spelling recognition tasks is upwards of 85% accuracy (much better than recall methods). In one such study, Yang and Perfetti (2006; as cited in Perfetti, 2007) used a lexical decision task after each of four word learning sessions. In this task, to assess learning of the spelling of new vocabulary words, participants had to choose whether a word was a real rare word that they were learning the meaning of in the study or if it was a foil of similar form. The three foil choices for real word target *hebetude* (which was to-be-learned over four sessions) were a word with high phonemic overlap *hebitude*, a word with lower phonemic overlap but high orthographic overlap *hebetide*, and a control word with low phonemic and orthographic overlap *hodilane* (i.e., only had
to share the same initial letter). Data from participants’ error rates in the lexical decision task demonstrated differences among foil types: foils with high phonemic overlap were chosen more often than foils with high orthographic overlap (but lower phonemic overlap). Both of these foil types were chosen more often than foils with low phonemic and orthographic overlap. Yang and Perfetti’s data suggest that choosing a foil with high phonemic overlap on a test of spelling recognition memory at higher rates than foils with high orthographic overlap would indicate that readers use newly generated phonological representations for novel words from silent reading to inform recognition of new word spellings.

To that point, previous studies using ERP (Ashby, 2010) and fixation time measures (e.g., Ashby & Clifton, 2005; Folk, 1999; Folk & Morris, 1995) have demonstrated that readers automatically activate and store phonological representations of words during silent reading. Recently, Brusnighan et al. (2014) provided fixation time evidence that readers automatically activate and store phonological representations for new vocabulary words from reading as well. Specifically, readers’ fixation durations on novel homophone words (i.e., a novel word skwosh shares a phonology with real word ‘squash’) were longer than on novel nonhomophone words like chenth (which does not share a phonology with any real word). Brusnighan et al. argued that reading times were longer for novel homophone words because when a word like skwosh is encountered, the meaning for squash is also activated and interferes with assigning a new meaning for the word, whereas this competition for meaning does not occur for a non-homophonic word like chenth. In adding to this evidence that readers activate and learn pronunciations for
new word spellings during silent reading, further post-reading tests of phonological learning are needed. Naming time has been used as a measure of lexical access. That is, naming times are faster for words than nonwords and faster for high frequency words than low frequency words. This demonstrates that naming a letter string occurs as a result of searching the lexicon for lexical entries (Forster & Chambers, 1973). Naming times for novel words would be slower than naming times for known words, but if readers begin to add entries for novel words to the mental lexicon upon initially encountering them from silent reading (and strengthening LQ with repeated exposure) then naming times may be expected to be faster for novel words that are encountered multiple times in a variety of contexts than novel words encountered once within a single reading session. If not, however, future research could explore how many exposures (and over what period of time) are needed before naming times for novel words begin to resemble naming times for known words that are matched in length and orthographic familiarity.

Effects of Contextual Constraint and Reading Exposure in Reading Novel Words

In reading of known words, it has been well documented that if a word is preceded by a semantically constraining context it is processed faster than if prior context were neutral (see Clifton, Staub, & Rayner, 2007 for a review). Applied to reading of novel words, researchers have established via fixation times and post-reading vocabulary test measures that readers infer new word meanings from context during reading (Brusnighan & Folk, 2012; Brusnighan et al., 2014; Chaffin et al, 2001; Williams & Morris, 2004). After an initial exposure to a novel word in an informative context from which a meaning inference was made during reading, readers likely begin to develop a
long term memory representation for the novel word (i.e., given context cues related to a novel word’s meaning prior to encountering it in a sentence, the novel word’s spelling may become preactivated versus when prior context is neutral). Thus, it is expected that over several exposures, new words preceded by informative contexts will be read faster than words preceded by neutral contexts.

Researchers have provided fixation time evidence of effects of number of exposures to novel words embedded in sentences. Several studies have provided evidence that when a known word is presented twice during a reading session, fixation times are shorter on the second presentation of the word (Binder & Morris, 1995; Raney & Rayner, 1995; Rayner, Pacht, & Duffy, 1994). Similarly, Joseph, Wonacott, Forbes, and Nation (2014) found that over several reading session, novel words are processed faster upon the second presentation. Extending this, it is expected that within a single reading session novel words will be processed faster with each reading exposure.

**Post-reading Tests of Semantic Learning from Incidental Exposure during Silent Reading**

Semantic learning from incidental exposure during reading is often measured using post-reading multiple choice vocabulary tests (Brusnighan & Folk, 2012; Brusnighan et al., 2014; Williams & Morris, 2004). However, many of these tests have not required differentiation between items learned in the learning task but rather a distinction between a meaning seen in the learning task and an unseen meaning. Herman, Anderson, Pearson, and Nagy (1987) argued that acquisition of word knowledge proceeds in small increments. Nagy and Anderson (1984) estimated that of the majority...
of word meanings learned per year, 5-12% of these word meanings are learned from a single reading exposure. Herman et al. (1987) suggested that these initial encounters with novel words in context may lead to only partial understanding of word meanings. Researchers have contended that the semantic knowledge resulting from initial encounters with novel words in context, while fragile, is pliable in that it can be strengthened with subsequent exposures in a variety of contexts (Bolger et al., 2008; Jenkins, Stein, & Wysocki, 1984; van Daalen-Kapteijns, Elshout-Mohr, & de Glopper, 2001). Thus, from one encounter with a new word, readers may only get a sense for what a word means, and then on a multiple choice test would only be able to select a correct meaning for the word if there are no distraction choices. However, if the test requires differentiation between multiple choices with similar meanings to the target word or meanings for other target words from a word learning episode, a sense for what a word means would be insufficient to do well on the test for meaning.

Williams and Morris (2004) monitored readers’ eye movements as they read sentences containing novel and known target words embedded in sentences with contexts that were constraining as to the intended meaning of the target words. After the reading session, participants completed a surprise multiple choice test for the meanings of the novel words from the reading session. For example, participants read the sentence:

Sheila selected the masdor for her bouquet.

Then, on the surprise test, they received two choices for what a ‘masdor’ meant:

A: flower
B: jewel
In the reading session, the word “flower” was not explicitly used, thus readers were left to infer that masdors are flowers because flowers go in bouquets. However, there were no sentences that mentioned “jewel” explicitly or that were related to jewels in the way that bouquets are related to flowers. In this way, all that is required at the time of the multiple-choice test is recognition of some episodic information from the learning task. That is, as a participant, all one needs to remember is that they did not read about a jewel, so they can simply choose the meaning that was familiar from the reading session, whether or not they had associated it with the specific orthography being tested. In this study, accuracy on the meaning recognition vocabulary test was 62% which was better than chance accuracy of 50%. Brusnighan & Folk (2012) used the same two-choice test method and accuracy was 92%, still better than chance of 50%. In one way of addressing the issue that readers in previous studies have not had to differentiate between meanings inferred from the reading session on vocabulary tests, Brusnighan et al. (2014) gave readers a test with four choices, one of which was a distractor choice and two of which were unrelated choices that were not read about in the reading session. The semantically related distractors were from something else read about in the sentences from the reading session. For example, readers saw the sentence pair:

The collapse of the coal mine shaft killed a blarth yesterday.

Everyone in town is mourning the worker's tragic death.

Then on the surprise test, they received four choices for what a ‘blarth’ meant:

A) concierge
B) coal gatherer
Thus, the choice between coal gatherer and elevator shaft requires some differentiation between aspects of the same episodic cue. However, readers did not have to differentiate between the meaning of a blarth and a masdor, for instance. Readers were 58% accurate (better than chance of 25%), chose the distractor choice 24% of the time, and chose each of the other two unrelated choices a combined total of 18% of the time. This is evidence that readers can learn meanings of new words from one exposure and remember them after the reading session. However, to test the specificity of semantic learning, evidence is needed from a test in which readers read about a blarth and a masdor, for instance, and have to choose whether blarths and masdors are each flowers or coal gatherers.

**The Current Study**

The current study investigates the effect of contextual constraint on reading times and post-reading memory for novel monomorphemic words encountered during silent reading. In doing so, the following research questions will be investigated.

*Do adult readers demonstrate more orthographic and phonological learning for novel words seen in informative contexts than neutral contexts?* Based on previous ERP findings that learning of a word’s phonology and learning of a word’s orthography both benefit from learning of that word’s meaning (Balass et al., 2010), it is expected that adult readers will demonstrate more orthographic and phonological learning on post-reading tests of learning for novel words seen in informative contexts than neutral
contexts. However, since the ERP study was an explicit word learning study, it is not clear that the results will apply to an incidental word learning task.

*Is one encounter with a novel word in context sufficient with no prior knowledge of the word to determine a meaning for the word on a multiple choice test or are more exposures needed? Based on prior research (Brusnighan & Folk, 2012; Brusnighan et al., 2014; Williams & Morris, 2004) suggesting that readers retain meanings for novel words from one reading exposure (albeit whose tests for meaning did not require differentiation between meanings inferred during the reading session), it is expected that readers will again demonstrate above chance accuracy for novel word meanings from a single reading encounter on a new vocabulary test in which a novel word’s meaning must be paired with its spelling. Further, consistent with prior research demonstrating that readers build LQ over repeated exposures in a variety of contexts (Frishkoff et al., 2008; 2009; 2010), it is expected that readers will demonstrate more semantic knowledge from several varied contexts for new words than from a single context for a new word.

*Is one reading exposure sufficient for readers to demonstrate learning of form (orthography/phonology) or are more exposures are needed? Based on prior research using spelling recognition tests in intentional word learning sessions (Yang & Perfetti, 2006), it is expected that readers will be above chance accuracy for recognizing novel word spellings from one exposure in an incidental word learning task as well. With regard to demonstrating learning of phonology using a naming time task, however, no prior research offers a prediction for whether naming times would be faster for words seen six times during reading vs. words seen once or whether more reading exposures in
a single session (or over several sessions) are needed in order to demonstrate learning via this task.
EXPERIMENT 1

Experiment 1 was designed to investigate (1) the effects of semantic learning on word form learning from reading, (2) the effects of contextual constraint in processing novel monomorphemic words in English during a single reading session, and (3) semantic, orthographic, and phonological learning from one reading encounter vs. several encounters in one sitting. Each prediction outlined above based on these research questions was tested by the crossing of two factors: prior sentence contextual constraint (neutral vs. informative) and number of reading exposures to a novel word during the reading session (one vs. six). In Experiment 1, readers’ eye movements were monitored during reading and memory for the spellings, pronunciations, and meanings of those novel words was tested after the reading session. Fixation times were used to measure effects of contextual constraint on processing of novel words during silent reading. Memory for the pronunciations of novel words was investigated by collecting naming times and judging pronunciation accuracy. Memory for the spellings and meanings of novel words was tested using new tests of meaning and spelling recognition. Again, these post-reading tests of word learning were designed to address confounds of previous tests used in word learning studies. Specifically, these tests are not confounded with episodic memory for the task material. Critically, conditional analyses based on meaning recognition data were predicted to qualify effects of sentence contextual constraint in spelling and naming time data. That is, to provide the best test of the notion that when a form is tied to meaning during reading memory for form is better after the reading
session, analyses of pronunciation and spelling accuracy as a function of vocabulary accuracy were planned. This will answer the question: Does accurately recognizing the meaning of a novel word increase accuracy of pronouncing the novel word or recognizing the correct spelling of the novel word?

Method

Participants. Fifty-four Kent State University students participated for course credit\(^1\). Participants were native English speakers, had uncorrected vision, and reported no reading disabilities. Participants agreed to release their ACT and/or SAT scores via informed consent for the purpose of ensuring that the assignment to counterbalancing conditions resulted in equivalent groups. The mean ACT/SAT verbal percentile rank (based on age) was 59% (SD = 23) across conditions. Each of 4 counterbalancing condition means fell within two-fifths of a standard deviation from the overall mean.

Materials.

Target words. Target words consisted of 32 novel, single-syllable, monomorphemic words ranging from 5-8 characters in length, and are listed in the Target Word column of Appendix B (e.g., *blaph, flurst*). Target words were operationally defined as “novel” in that they were at least 2 characters different from any known English word, according to a traditional dictionary and an internet search. The novel

\(^1\) An *a priori* power analysis based on pilot study results was run to determine appropriate sample size needed for Experiment 1. The analysis revealed that at least forty-eight participants would be needed for an effect size $F$ of .25.
words were created such that only orthographically existing onsets, orthographically existing bodies, and legal bigrams in English were used. Target words were formed by selecting 32 different existing consonant clusters (e.g., bl-, fl-) as word onsets and combining them with 32 different existing spelling bodies (e.g., -aph, -urst). The result was 32 CVC (consonant-vowel-consonant) structure words. The spelling bodies were selected using the Ziegler, Stone, and Jacobs (1997) database such that all words were feedforward consistent (i.e., the spelling body mapped onto only one pronunciation) but feedback inconsistent (i.e., the pronunciation mapped onto more than one spelling body; see Lacruz & Folk (2004) for a review). In this way, there was only one plausible way to pronounce the spellings blaph or flurst but there was more than one plausible spelling for each pronunciation. The reason for creating words that followed this pattern was to reduce variability in the naming time task by restricting possible pronunciations. Finally, the frequency of phoneme-to-grapheme correspondences was controlled such that half of the target words’ spelling bodies contained the more likely correspondence (-ate for /eɪt/) and half the less likely (-aïght for /eɪt/) so that frequency of spelling (a measure of the regularity of the spellings) would not be a clue to the correct spelling when choosing between the target word and a homophone lure on the spelling recognition test.

**Sentence contexts.** Target words were embedded in sentence frames that were either neutral or informative as to the intended meaning of the target words (see Appendix A for complete stimuli). Stimuli were single sentences that fit on a computer screen on a single line of text. Experimental sentence frames were created such that a context region preceded the novel target word. For informative sentences, the context
region contained semantic cues from which to derive a meaning for the novel word but a
definition for the novel word was not explicitly given. Neutral sentence contexts were
written such that virtually any known English noun could be inserted in the target word
position and make sense. The target words in the sentences were followed by a short,
normal wrap-up region that completed each sentence. Filler sentences were also created,
containing no novel words in order to preserve the naturalistic validity of a normal
reading environment (i.e., approximately 40% of sentences presented during reading
contained no novel words).

There were four experimental conditions created by the crossing of two factors: sentence contextual constraint (informative or neutral sentences); and number of exposures/number of sentences in which novel words appeared (one or six sentences). In all, there were eight words to appear one time each in an informative sentence context (8

2 Sentence frames were normed for sentence contextual constraint. First, seventy participants rated each sentence frame (with a blank in the target word position) on how informative or neutral it was as to the meaning of the target word on a 5-point likert scale (1 = neutral and 5 = very informative). Mean informativeness ratings were 3.69 (SD = 0.31) for informative sentences and 1.97 (SD = 0.22) for neutral sentences. Next, a separate group of seventy-one participants were presented with each sentence context with a blank space in place of the target word and asked to complete the sentence with the first word or phrase that came to mind. Three raters who were naïve to the conditions of the study determined that all sentence frames fell within a priori cutoffs and should be included in the study. Inter-rater agreement was 91%. For neutral sentences, the most common response across sentences constituted 14% (SD = 6%) of responses. For informative sentences, 84% (SD = 14%) of responses were related to the category to which the intended definition belonged.
sentences), eight words to appear six times each embedded in different informative sentence contexts (48 sentences), eight words to appear once in neutral sentence contexts (8 sentences) and eight words to appear six times in different neutral sentence contexts (48 sentences). Thus, there were eight items per experimental condition resulting in 112 experimental sentences.

**Spelling recognition test.** In this task, there were 4 possible response types (see Appendix B). The target word spellings were the correct choice in that they matched the spellings seen during the reading session (Target Word column). Three alternate incorrect choices were created for each target word for use in this new task. First, an alternate way to spell the only plausible pronunciation of each target word was given (a homophone of the target word (Homophone Lure column)). Half of the homophone lures contained the more likely (i.e., higher frequency spelling) phoneme-grapheme correspondence spelling bodies when the less likely correspondences were used in target words, and vice versa (see Appendix C for spelling body frequencies). The result was that this choice has high phonemic overlap but low orthographic overlap with the target word.

Researchers have shown that a *serial position effect* occurs in spelling such that people most often misspell the middle of words (Jones, 2013; Jones, Folk, & Rapp, 2009). Jones et al. (2009) suggest that this occurs because there is variability in the representational strength across elements of an orthographic representation (this variability is known as *orthographic texture*). Thus, since readers are more likely to make spelling errors on medial letters of words than initial or final letters of words, the second alternate incorrect choice for each target word on the spelling test (Transposition Lure
column) was created by transposing medial letters but preserving the initial and final letter identities. Transposing medial letters was done in an orthographically legal way, and if this was not possible, one medial letter was deleted or substituted for a legal letter in that position. The result was that this choice on the spelling test has higher orthographic overlap with the target word but lower phonemic overlap with the target word than the homophone lure words. Finally, a third incorrect choice was created by transposing medial letters (or substituting/deleting a letter) of the homophone lure words (Transposed Homophone column). The result was that this choice on the spelling test had lower orthographic overlap with the target words than the other two lure choices on the test.

**Meaning recognition vocabulary test.** In this task, for each of 32 target words there were 16 choices (see Appendix D for examples). For the 16 words seen by participants in informative contexts, choices were a correct definition and 15 incorrect definitions (the incorrect definitions for one word each corresponded to one of the other 15 novel words seen in informative contexts). In this way, each of those 15 incorrect definitions for one word was a correct definition for one other word in the Experiment. For example, participants had to correctly identify that a *blaph* was a type of musical instrument and not one of the other 15 definitions that corresponded to other words seen during reading. Thus, chance accuracy was 6.25%. For the 16 words seen in neutral text, 16 definitions that did not relate to any materials from the reading study were used as meaning choices. In this way, “accuracy” for defining words from neutral texts was arbitrarily and randomly assigned *a priori*. That is, each novel word was assigned a
meaning by the experimenter, which was then designated as the correct answer on the test.

**Procedure.** Participants were tested individually and were informed that they would complete several tasks involving reading and spelling. They were informed that during the reading session (their first task) they would see some new vocabulary words, but that they should still be able to comprehend what they were reading and that their goal during the reading session was, in fact, to comprehend everything they read.

Before testing, a bite bar was prepared for each participant to rest their mouth on during reading to minimize head movements. The eyetracking system was calibrated for each participant before reading any materials, taking approximately 5 minutes. A row of five target boxes was presented on the computer screen to start each trial so that the experimenter could ensure calibration of the eye tracking system throughout the task. To begin a trial, participants looked at the leftmost target box, at which time the experimenter presented a trial. This leftmost target box marked the position of the first letter of the sentences. Each participant was instructed to press a button when he or she finished reading a sentence. This buttonpress removed a sentence from the screen, replacing it with the row of five target boxes. Participants read six practice trials and answered one comprehension question to familiarize them with the procedures before testing. During reading of experimental materials, approximately 10% of the sentences were randomly followed by a “Yes/No” question to ensure that participants comprehended what they were reading. Comprehension questions only followed filler sentences and all participants scored at 90% or above on the comprehension questions (M
Participants were not told that the reading session would be followed by tests for knowledge of the materials presented during the reading session. In this way, experimental post-reading tests of word learning were a surprise. In addition to the 112 experimental sentences, there were 72 filler sentences. The selection of the target word condition for sentence frames was counterbalanced in a Latin Square design such that novel words seen only in neutral sentences by one participant were used as novel words only in informative sentences for the next participant. Sentence trials were randomized such that the items seen in six different sentences were spread out in the reading task.

Following the reading session, participants completed surprise tests of knowledge for the words seen during the reading session, using E-Prime and DirectRT software to record accuracy and button-press/voice onset latency for response times. The first post-reading task was a spelling recognition test completed using E-Prime. Participants were instructed to choose the correct spelling (i.e., the spelling used during the reading session) from four choices on each trial using a PST response box. Presentation order for each of the four choices within a trial was randomized and trial order was randomized.

Following the spelling recognition test, participants completed a naming time task using DirectRT software. Participants saw each of the target words presented in isolation in the center of the screen one-at-a-time and were asked to pronounce each target word as quickly and accurately as possible into a microphone. The participant’s voice was the key to move on to each next trial. Reaction time was recorded at the onset of each verbal response. Each word was presented after a fixation cross for 1 second and a 500ms blank
space, and each word disappeared as soon as a response was recorded. Presentation order for each of 32 novel words was randomized.

Following the naming time task, participants completed a multiple choice vocabulary test for meaning recognition using E-Prime, concluding the experiment. In this task, participants were instructed to select the best meaning for each of 32 novel words from 16 choices each using the A-P keys on the computer keyboard. Presentation order for each of 32 novel words was randomized.

**Apparatus.** Participants’ eye movements were monitored using a Fourward Technologies (Jameson, MO) Generation 6.1 Dual Purkinje eye movement monitoring system. Viewing was binocular and eye location was recorded from the right eye. The system was interfaced with a computer which controlled stimulus display and data storage. The sentences were presented on a VGA monitor with up to 72 characters per line of text. Characters were displayed in lower case with the exception of the first character of a sentence or a proper name. Participants were seated 31 inches from the monitor and four character spaces subtended 1° of visual angle. Eye position was sampled every millisecond.

All 3 post-reading tasks were administered using E-Prime 2.0 and DirectRT software on a computer used for stimulus display and data storage. The computer was equipped with a keyboard and PST response box used to record response times via buttonpress as well as a microphone used to record naming times via voice-key.

**Results 1a: Eyetracking**

Linear Mixed-Effects (lme) models were fitted to each reading measure using R
statistical software (version 2.13.0). A 2 (exposures: 1 v. 6) x 2 (context: informative v. neutral) design was used to explore effects of contextual constraint on novel word reading. Participants and items were included in each model as crossed random effects. The $p$-values were calculated via Markov chain Monte Carlo sampling (see Baayen, 2008 for a detailed model description).

For all fixation time measures, fixations shorter than 120 milliseconds or longer than 1000 milliseconds were excluded from the analyses. Fixation times occurring outside of this range are more often the byproduct of an eyelink or track loss than an extended or truncated fixation (Rayner, 1998). Approximately 3% of trials were lost due to the trimming procedure and trials lost were equally distributed among conditions. Length varied among items but was matched across conditions for the context regions in informative and neutral sentence frames. Zero durations (i.e., when a word was skipped) were excluded from gaze durations but included for rereading times and total reading times.

**Effects of contextual constraint in reading novel words.** To investigate effects of contextual constraint on reading novel words in either one or several reading exposures during the reading session, gaze duration, target rereading time, and total reading time on the target word measures were analyzed. Fixation time measures are reported in milliseconds. Gaze duration is a measure of initial reading time on the target word and is calculated as the sum of all fixations on the target word from when a reader first encounters the word until the reader leaves it. Target rereading time is a measure of reanalysis of the target word and is calculated as the sum of all fixations on the target
word from second enter to second exit. Total reading time on the target is a measure that includes both initial reading time and reanalysis of the target word and is calculated as the sum of all fixations on the target word from first enter to last exit.

In gaze durations, the analysis revealed a main effect of exposures (see Table 1 for significance values and Table 2 for means), indicating that readers processed novel words for significantly longer durations on sole encounters with novel words (in the 1 exposure condition) than on their sixth exposure to a novel word (in the 6 exposures condition). The analysis also revealed a main effect of contextual constraint such that readers spent less time reading novel words preceded by informative information from which to derive a meaning than when preceding text was neutral. The interaction between number of exposures and contextual constraint was significant, such that a larger difference between reading times in the one exposure and six exposures conditions was found for novel words in neutral sentences than in informative sentences.
Table 1

Mixed-effects Models for Target Word Processing in Experiment 1.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Fixed effect</th>
<th>b</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gaze Duration Target</strong></td>
<td>Exposures</td>
<td>-12.4</td>
<td>1.9</td>
<td>-6.42</td>
<td>.0001</td>
<td>****</td>
</tr>
<tr>
<td></td>
<td>Informative</td>
<td>-55.7</td>
<td>15.4</td>
<td>-3.63</td>
<td>.0003</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>8.6</td>
<td>2.8</td>
<td>3.12</td>
<td>.002</td>
<td>**</td>
</tr>
<tr>
<td><strong>Rereading Time Target</strong></td>
<td>Exposures</td>
<td>-15.3</td>
<td>1.8</td>
<td>-1.54</td>
<td>.118</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Informative</td>
<td>-25.8</td>
<td>10.7</td>
<td>-2.72</td>
<td>.007</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>4.4</td>
<td>3.0</td>
<td>1.46</td>
<td>.136</td>
<td></td>
</tr>
<tr>
<td><strong>Total Time Target</strong></td>
<td>Exposures</td>
<td>-15.9</td>
<td>3.4</td>
<td>-4.66</td>
<td>.0001</td>
<td>****</td>
</tr>
<tr>
<td></td>
<td>Informative</td>
<td>-72.0</td>
<td>26.9</td>
<td>-2.67</td>
<td>.008</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>9.7</td>
<td>4.8</td>
<td>2.00</td>
<td>.046</td>
<td>*</td>
</tr>
</tbody>
</table>

*Note.* Significance codes: ‟†” < .1, ‟*” < .05, ‟**” < .01, ‟***” < .001, ‟****” < .0001.

The fixation time pattern in total reading time (which also included rereading times) was similar to that found for gaze duration, as the analysis for total reading time revealed significant main effects of number of exposures and sentence contextual constraint and an interaction (see Table 2 for means). For rereading times on the target word, the analysis revealed a main effect of sentence contextual constraint such that, again, novel words preceded by neutral contexts were reread for longer durations than novel words preceded by informative contexts. However, an effect of number of exposures and interaction between exposures and contextual constraint did not approach significance.
Table 2

*Means for Reading Times on the Target Word by Condition in Experiment 1.*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze Duration</th>
<th>Rereading Time</th>
<th>Total Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-Informative</td>
<td>359 (15)</td>
<td>324 (24)</td>
<td>429 (22)</td>
</tr>
<tr>
<td>One-Neutral</td>
<td>405 (19)</td>
<td>375 (24)</td>
<td>491 (30)</td>
</tr>
<tr>
<td>Sixth-Informative</td>
<td>321 (11)</td>
<td>313 (29)</td>
<td>345 (11)</td>
</tr>
<tr>
<td>Sixth-Neutral</td>
<td>337 (11)</td>
<td>366 (72)</td>
<td>343 (12)</td>
</tr>
</tbody>
</table>

*Note.* Reading times are given in milliseconds. Standard errors are in parentheses.

Means from each trial within a set in the six exposures conditions for informative and neutral conditions (see Table 3) revealed that over trials in the six exposures condition there was a repetition effect such that readers required less total reading time on the novel word with each exposure ($b = -38.802, SE = 2.603, t = -14.91, p < .0001, R^2 = .88$).
Table 3

*Total Reading Times on Target Words for Repeated Items by Trial Order in Experiment 1.*

<table>
<thead>
<tr>
<th>Measure/Condition</th>
<th>Statistic</th>
<th>Trial Order</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Total Time Target</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>M</td>
<td>548</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>13</td>
</tr>
<tr>
<td>Informative</td>
<td>M</td>
<td>480</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>M</td>
<td>514</td>
</tr>
</tbody>
</table>

*Note.* Reading times are given in milliseconds. *SE* = standard error of the mean.

**Results 1b: Post-reading Tests of Word Learning**

Lme models were again fitted to reaction time measures (e.g., naming times) in a 2 (context: informative v. neutral) x 2 (exposures: 1 v. 6) design. For measures in which the dependent variable was a percentage (e.g., accuracy on the vocabulary test) a logistic model was fitted to the data, again in a 2 x 2 design, by specifying that it followed a binary distribution (amounting to mixed-effects logistic regression). Additionally, since there were four categorical choices on the spelling recognition test, each response type was analyzed within a cumulative link mixed model (clmm), again in a 2 x 2 design. Specifically, the clmm measures whether context or exposure impacted the frequency with which participants chose a spelling within each of the four response types on the
spelling recognition test (rather than treating responses as merely accurate or inaccurate in a binary logistic regression, see Christensen, 2012 for further description).

**Meaning recognition vocabulary test results.** In the analysis for accuracy on the vocabulary test, the main effect of exposures did not approach significance, but the main effect of contextual constraint did approach significance and the interaction reached significance (see Table 4 for significance values). A review of the means from this measure (Table 5) revealed that accuracy was low (near chance levels) across conditions. Therefore, a statistically significant interaction in this measure is not likely to be meaningful. However, support for the prediction that meanings for words seen six times would be recognized with higher accuracy than for words seen once and that words seen in informative contexts would be recognized above chance (while words seen in neutral contexts would not) came from follow-up comparisons between each of three conditions (neutral 1 and neutral 6 conditions were combined as no meaning was given) and between each of these conditions and chance accuracy levels. Only the six exposures-informative condition was marginally better than chance levels of accuracy ($t(53) = 1.82, p = .074$) and only the six exposures-informative condition significantly differed from the one exposure-informative condition ($t(53) = -2.81, p = .0059$). The six exposures-informative condition was not significantly better than both neutral conditions ($t(53) = -1.59, p = .116$) but, importantly, the one exposure-informative and neutral conditions did not differ from each other ($t(53) = 0.82, p = .414$). Also, as predicted, the fact that word
meanings in the neutral sentence condition were not recognized above chance replicates previous findings (Brusnighan & Folk, 2012), as no meaning was given.

Table 4

Mixed-effects Models for the Post-reading Tests of Learning in Experiment 1.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Fixed effect</th>
<th>b</th>
<th>SE</th>
<th>t or z</th>
<th>P</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary Accuracy</td>
<td>Exposures</td>
<td>-0.07</td>
<td>0.05</td>
<td>-1.32</td>
<td>.187</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Informative</td>
<td>-0.70</td>
<td>0.42</td>
<td>-1.67</td>
<td>.096</td>
<td>†</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>0.19</td>
<td>0.09</td>
<td>2.02</td>
<td>.043</td>
<td>*</td>
</tr>
<tr>
<td>Spelling Recognition Test</td>
<td>Exposures</td>
<td>0.11</td>
<td>0.04</td>
<td>2.42</td>
<td>.015</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Informative</td>
<td>-0.13</td>
<td>0.26</td>
<td>-0.51</td>
<td>.614</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>-0.01</td>
<td>0.06</td>
<td>-0.16</td>
<td>.874</td>
<td></td>
</tr>
<tr>
<td>Naming Times</td>
<td>Exposures</td>
<td>-7.64</td>
<td>5.07</td>
<td>-1.51</td>
<td>.132</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Informative</td>
<td>-36.6</td>
<td>30.8</td>
<td>-1.19</td>
<td>.235</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>10.4</td>
<td>7.16</td>
<td>1.45</td>
<td>.146</td>
<td></td>
</tr>
<tr>
<td>Pronunciation Accuracy</td>
<td>Exposures</td>
<td>0.01</td>
<td>0.06</td>
<td>0.13</td>
<td>.898</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Informative</td>
<td>0.19</td>
<td>0.38</td>
<td>0.49</td>
<td>.623</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>-0.05</td>
<td>0.09</td>
<td>-0.60</td>
<td>.547</td>
<td></td>
</tr>
</tbody>
</table>

Note. Significance codes: † < .1, * < .05, ** < .01, *** < .001, **** < .0001.

Reaction time measures use t-values and logistic/clmm models use z-values.

There is no actual right answer for the novel words seen in neutral sentences, so the experimenter arbitrarily as a comparison control picked one. All choices should be chosen equally as often (a manipulation check for the materials to ensure the sentences were all equally neutral).
Table 5

Vocabulary Test Accuracy by Number of Exposures in Experiment 1.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral (no meaning)</td>
<td>6.7 (0.3)</td>
</tr>
<tr>
<td>Informative-One exposure</td>
<td>4.4 (0.1)</td>
</tr>
<tr>
<td>Informative-Six exposures</td>
<td>9.3 (0.4)</td>
</tr>
</tbody>
</table>

Note. Accuracy is in percentages. Standard errors are in parentheses. Chance levels of accuracy were 6.25%.

**Spelling recognition test results.** For spelling recognition accuracy, the analysis revealed a main effect of exposure with no effect of contextual constraint and no interaction (see Table 4 for significance values). The main effect of exposure was consistent with predictions based on the LQH, such that word spellings seen over six contexts were recognized with higher accuracy than words seen once during reading. Although the effect of sentence contextual constraint was not significant, means for this measure between context types (Table 6) were in the opposite direction of what was predicted such that word spellings seen in neutral contexts were recognized with higher accuracy than words seen in informative contexts. The implications of this pattern will be addressed in Experiment 2.
Table 6

_Means for Spelling Recognition Accuracy in Experiment 1._

<table>
<thead>
<tr>
<th>Condition</th>
<th>Spelling Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-Informative</td>
<td>33.1 (2.9)</td>
</tr>
<tr>
<td>One-Neutral</td>
<td>38.9 (2.5)</td>
</tr>
<tr>
<td>Six-Informative</td>
<td>47.5 (2.7)</td>
</tr>
<tr>
<td>Six-Neutral</td>
<td>53.9 (2.8)</td>
</tr>
</tbody>
</table>

*Note.* Accuracy is in percentages. Standard errors are in parentheses.

**Response comparisons.** Overall, participants’ accuracy in identifying the correct spellings for novel words was better following six exposures to novel words than following one exposure to novel words, yet there were not differences in response proportions between each of the four response options for words seen in informative versus neutral contexts (see Table 7 for means by exposure condition for each of four response types). Collapsing across exposure conditions, among lures on the spelling test, participants were more likely to choose a Homophone Lure than either of the transposed options (*t*(55) = 10.15, *p* < .0001), with no difference in frequency of choosing between transposed options (*t* < 1).
Table 7

Means for Spelling Recognition Test Responses by Exposure Condition in Experiment 1.

<table>
<thead>
<tr>
<th>Response Type</th>
<th>Example Item</th>
<th>One Exposure</th>
<th>Six Exposures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Spelling</td>
<td>blaph</td>
<td>36.2 (1.9)</td>
<td>50.1 (2.1)</td>
</tr>
<tr>
<td>Homophone Lure</td>
<td>blaff</td>
<td>34.7 (2.0)</td>
<td>26.9 (1.5)</td>
</tr>
<tr>
<td>Transposition Lure</td>
<td>balph</td>
<td>14.3 (1.2)</td>
<td>12.2 (1.2)</td>
</tr>
<tr>
<td>Transposed Homophone</td>
<td>balff</td>
<td>14.8 (1.4)</td>
<td>10.8 (1.3)</td>
</tr>
</tbody>
</table>

*Note.* Means are in the percentage of the total responses belonging to each response type (adding up to 100%). Standard errors are in parentheses.

**Naming time task results.** Participants’ pronunciations were coded for accuracy on the basis of two criteria. If the participant’s pronunciation was a plausible pronunciation for the novel word, the response was coded as correct. However, if the participant produced a pronunciation that was not a phonologically plausible interpretation of the letter string, stuttered, hesitated, or corrected their pronunciation, the response was coded as incorrect. Naming times that were two standard deviations above or below a participant’s mean were also removed from analyses. Data loss was equally distributed across conditions, 17.5% of trials were removed from analyses due to being incorrect pronunciations, and 2.1% of trials were removed from analyses due to being more than two standard deviations outside of participants’ mean reaction times.

It was expected that words encountered six times would be named faster and with higher accuracy than words encountered once and that words seen in informative contexts would be named faster and with higher accuracy than words seen in neutral texts, but
neither naming times nor pronunciation accuracies in Experiment 1 differed between conditions (see Table 7 for significance values and Table 8 for means).

Table 8

*Mean Naming Times and Pronunciation Accuracies in Experiment 1.*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Naming Time</th>
<th>Pronunciation Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-Informative</td>
<td>940 (44)</td>
<td>83.8 (2.1)</td>
</tr>
<tr>
<td>One-Neutral</td>
<td>965 (47)</td>
<td>82.2 (1.5)</td>
</tr>
<tr>
<td>Six-Informative</td>
<td>940 (43)</td>
<td>81.3 (1.2)</td>
</tr>
<tr>
<td>Six-Neutral</td>
<td>924 (41)</td>
<td>82.9 (1.3)</td>
</tr>
</tbody>
</table>

*Note.* Naming times are in milliseconds. Accuracy is in percentages. Standard errors are in parentheses.

**Discussion**

In eyetracking measures in Experiment 1, analyses revealed effects of sentence contextual constraint and number of reading exposures on novel word reading times, in a to-be-expected direction: reading times were faster on a sixth exposure to a new word than on a single encounter with a new word, and in the six exposures condition readers required less total reading time on the novel word with each exposure. This finding with novel words read within a single reading session is consistent with effects of repetition with familiar words (Binder & Morris, 1995; Raney & Rayner, 1995; Rayner et al., 1994) and with novel words between reading sessions (Joseph et al., 2014), such that fixation times are shorter on the second presentation of a word. Also, novel words preceded by informative contexts required less total viewing time. This second finding may be an
intuitive one: readers spend more time trying to infer a meaning for a novel word (or integrate that word into the sentence context) when no meaning cues are available. Also, in the case of novel words seen in informative contexts, there is likely some preactivation of an emerging lexical representation with each new context consistent with a novel word’s meaning (after the first encounter in which a meaning was inferred) which provides a reading time advantage relative to reading of a novel word preceded by neutral context.

The critical manipulation in Experiment 1 was that of contextual constraint as a way to investigate learning of word form with and without learning of word meaning. However, it was not possible to investigate the relationship between semantic learning and the learning of word form via conditional analyses because (1) accuracy was too low on the vocabulary test and (2) there was too much variability in participants’ naming time latencies. These issues were addressed in designing Experiment 2.

Webb (2008) concluded that spelling accuracy for novel words is impacted by number of reading exposures rather than context quality. In Experiment 1, spelling accuracy for novel words was significantly impacted by number of reading exposures, but it is still unclear whether context quality plays a role in spelling accuracy for novel words. To that point, the spelling test in Experiment 1 revealed a surprising pattern. Although the pattern was not statistically significant, participants showed better spelling performance for novel words seen in neutral than informative contexts. One possible reason for this pattern is that throughout measures of reading and rereading of the novel words, a main effect of sentence contextual constraint was found such that readers spent
significantly more time viewing novel word spellings in neutral texts than those in informative contexts. Thus, this extra viewing time for word spellings in neutral sentences could have led to higher recognition accuracy for these spellings on the spelling test administered directly after the reading session. Thus, in Experiment 2, rereading was disallowed in an attempt to test whether rereading times can account for differences in spelling accuracy (i.e., if the experimenter can control for reading times such that they are equal between groups and spelling accuracies are then equal between groups, accuracy differences are accounted for by reading times).

Experiment 1 did not yield compelling evidence from the new test of meaning recognition that readers are actually able to identify the spelling that corresponds with a specific definition for a novel word seen during reading. Because accuracy was so low on the vocabulary test, this test could have been too difficult because there were too many novel words for participants to learn in one sitting. Previous studies’ post-reading vocabulary tests have only contained 8 (Williams & Morris, 2004) or 16 (Brusnighan et al., 2014) to-be-learned items. The vocabulary test in Experiment 1 contained 32 to-be-learned items. Thus, in Experiment 2 the number of to-be-learned words was reduced to 16 to-be-learned items by removing the one exposure conditions. Removing the one exposure condition was possible because effects of number of reading exposure were demonstrated in fixation time measures and in spelling recognition accuracy.

Of predicted outcomes, naming time differences and pronunciation accuracy differences were not observed. This issue could be explained by variability in each participant’s response times in Experiment 1. By removing the half of the novel word
stimuli with the most pronunciation errors and adding a practice naming time session in Experiment 2, variability in naming times should be reduced.

Brusnighan et al. (2014) demonstrated that readers generate phonological representations for novel words upon their first incidental encounter with those novel words during silent reading. Among errors on the post-reading spelling test in Experiment 1 of the current study, results extended the pattern found by Yang and Perfetti (2006) such that homophone lures were chosen more often than transposed versions of the correct spellings. This pattern suggests that readers use newly generated phonological representations for novel word spellings encountered during reading in informing decisions of how to spell those novel words after the reading session. Alternatively, cues from the orthographic familiarity of spelling choices could have also lead readers to pick a homophone option more often than a transposed option if homophone spellings were more orthographically regular. White (2008) suggested that token bigram frequencies provide an appropriate measure of orthographic familiarity. To explore this alternative in the current study, bigram frequencies were calculated for each spelling in four choice conditions on the spelling test to determine if the reason readers chose the homophone spellings more often than the transposed spellings is that they had higher bigram frequencies than transposed choices. Bigram frequencies did not differ between these two choice conditions ($t(31) = 0.07, p = .947$) or, for that matter, between any of the four conditions ($\chi^2 = 2.94, p = .401$; see Table 9 for means).
Table 9

Means for Token Bigram Frequencies of Spelling Recognition Test Responses in Experiment 1.

<table>
<thead>
<tr>
<th>Response Type</th>
<th>Example Item</th>
<th>Bigram Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Spelling</td>
<td>blaph</td>
<td>575 (77)</td>
</tr>
<tr>
<td>Homophone Lure</td>
<td>bluff</td>
<td>671 (112)</td>
</tr>
<tr>
<td>Transposition Lure</td>
<td>balph</td>
<td>657 (165)</td>
</tr>
<tr>
<td>Transposed Homophone</td>
<td>balff</td>
<td>801 (166)</td>
</tr>
</tbody>
</table>

*Note.* Token bigram frequencies were calculated using the CELEX lexical database (Baayen, Piepenbrock, & Van Rijn, 1995). Standard errors are in parentheses.

Means for token bigram frequencies across conditions further suggest that readers use newly generated phonological representations for recently encountered novel words as a cue to spelling. That is, it does not seem to be the case that orthographic familiarity was a reliable cue to spelling recognition (and was therefore not the reason that readers more often chose a homophone spelling than a transposed spelling). Thus, when readers did not recognize the correct spelling for a novel word encountered during reading, their errors were most often based on a plausible pronunciation for the novel spelling rather than the specific letters used in that spelling during reading.
EXPERIMENT 2

Experiment 2 was designed primarily to address the low accuracy rates on the vocabulary test in Experiment 1. That is, vocabulary test accuracies were either not significantly different from chance because participants cannot learn the meanings of novel words (indicating an issue with previous studies’ vocabulary test methods) in so few exposures when eliminating episodic memory cues or because there was too much information to be learned from one reading session (indicating that the vocabulary learning requirements were too difficult). To investigate which issue affected vocabulary test accuracies, the amount of information to be learned in Experiment 2 was halved from Experiment 1 to better reflect the amount of learning required in other studies of incidental vocabulary learning (e.g., 16 novel words). Next, the results of Experiment 1 indicated that differences in reading times (higher reading times for neutral sentences) could have been responsible for spelling accuracy differences, in particular a trend for higher accuracy spelling for words in neutral sentences. Thus, in Experiment 2, a phrase-by-phrase moving-window paradigm for self-paced reading was implemented to disallow rereading. The purpose of this was to determine whether post-reading spelling accuracy rates in Experiment 1 can be explained, at least in-part, by target word rereading times. In this way, if reading times are not different on novel words by sentence context condition, it could be ruled out that reading times explain why spelling accuracy would be higher for a word seen in a neutral context.
Consistent with prior work (Herman et al., 1987; Perfetti, 2007; Perfetti & Hart, 2001; 2002; Yang & Perfetti, 2006) analyses of Experiment 1 spelling recognition and vocabulary test data (and numerical trends in naming time data) indicated that readers develop representations of form and meaning over several varied exposures during reading. Since the manipulation of one exposure versus six exposures supported predictions in Experiment 1, using only a six exposures condition in Experiment 2 allowed for the amount of information to be learned from reading to be halved.

Similar to the issue that low vocabulary test accuracies presented, either naming times did not differ between conditions because participants were not building stability in representations of phonological form over repeated exposures from reading (and not tying phonological form to meaning during reading) or there was too much variability in participants’ responses to detect differences between conditions statistically (i.e., there were too many errors and because there was not a practice session to familiarize participants with how quickly they were expected to name words on a speeded naming time task, the range of reaction times across participants was too wide). To investigate which issue affected naming time data, half of the items from Experiment 1 were selected based on the fewest number of pronunciation errors to reduce variability in naming time data. Also, using the removed novel words as practice session words was expected to further reduce variability in naming time data.

In reading times on the novel words, repetition effects over six trials were again expected, replicating Experiment 1. It was further expected that in a self-paced reading task, since rereading is disallowed, reading times will not differ as a function of sentence
contextual constraint. Because of this, it was expected that spelling recognition accuracy would not be higher for words seen in neutral sentences than in informative contexts. If any effects of contextual constraint occur, word spellings seen in informative sentences would be recognized with higher accuracy than word spellings seen in neutral sentences because they are theoretically higher quality spelling representations. However, as Webb (2008) suggested, it may be number of exposures and not context quality for the sentences in which the new spellings were embedded during reading that impacts accuracy of orthographic knowledge after the reading session. If this is the case, spelling recognition accuracy should be the same for words seen in informative and neutral contexts.

It was expected that reducing the number of items to be learned from reading would increase post-reading meaning and spelling accuracies, and that vocabulary test accuracy in the informative condition where there was actually a definition to be learned would be significantly higher than chance levels from six reading exposures. In naming time measures, words seen in informative sentences should be named faster and with fewer errors than words seen in neutral sentences. Finally, conditional analyses based on meaning recognition data were again predicted to qualify effects of sentence contextual constraint in spelling and naming time data. That is, conditional analyses provide the best test of the notion that when a form is tied to meaning during reading, as reflected in a correct response on the vocabulary test, memory for form is better after the reading session.

Method
**Participants.** Fifty-two Kent State University students participated for course credit\(^4\). Participants were native English speakers and reported no reading disabilities. None had participated in Experiment 1. Participants’ ACT and/or SAT scores were again released for this sample to ensure that assignment to counterbalancing conditions resulted in equivalent groups. The mean ACT/SAT verbal percentile rank for each of two counterbalancing conditions was 65% \((SD = 20)\) and 61% \((SD = 23)\), respectively.

**Materials.** Sentence materials used in Experiment 2 were the same as those used in Experiment 1 minus the “1 exposure informative condition” and “1 exposure neutral condition” sections of Appendix A. There were 96 experimental sentences and 20 filler sentences. These filler sentences did not contain any novel words. Thus, in Experiment 2, there were 8 novel words to appear in 6 different informative contexts and 8 novel words to appear in 6 different neutral sentences (see Appendix E, Target Word column for words used in Experiment 2).

There were two conditions created from two levels of one factor: sentence contextual constraint (neutral or informative sentences). Selection of the target word condition for sentence frames was counterbalanced in a Latin Square design such that novel words seen only in neutral sentences by one participant were used as novel words only in informative sentences for the next participant. Sentence trials were randomized such that each of the six different sentences in which a novel word appeared were spread

\(^4\) An *a priori* power analysis based on Experiment 1 results was run to determine appropriate sample size needed for Experiment 2. The analysis revealed that at least forty-eight participants would be needed for an effect size \(F\) of .25.
out in the reading task. Spelling recognition test choices for Experiment 2 are in Appendix E. Example vocabulary test examples for Experiment 2 are in Appendix F.

**Procedure.** Procedures were the same in Experiment 2 as in Experiment 1, with the exception that reading trials were presented using a phrase-by-phrase moving window task in E-Prime. In this task, sentences were divided into three phrases and these phrases were presented one-phrase-at-a-time such that the participant pressed a button to remove a previous phrase and bring up the next phrase in isolation. The first phrase that was presented was the context that preceded the novel target words, followed by the novel target words, followed by a short, neutral sentence wrap-up region. In this manner, when the participant moved from one phrase to the next, rereading of text was not allowed. Participants controlled the onset of each reading trial. Participants pressed a button to begin each reading trial and pressed a button again when they finished reading each of the three phrases for each sentence. Pressing a button when they finished reading the third phrase removed the third phrase from the screen and ended the trial. Participants read four practice trials and answered one comprehension question to familiarize them with the procedures. During reading of experimental materials, approximately 18% of the sentences read were randomly followed by a comprehension question. Comprehension questions only followed filler sentences. All participants scored at or above 90% accuracy on the comprehension questions ($M = 96\%$). Following the reading session, participants completed surprise tests of knowledge for the words seen during the reading session using the same procedures as in Experiment 1 with the exception that participants completed a practice naming time task (using the 16 words that were unused from
Experiment 1 that served as one control measure for naming times) and then completed the experimental naming time task (to further reduce variability in the experimental naming times) using DirectRT software.

**Apparatus.** Both the self-paced reading and post-reading tasks were administered using E-Prime 2.0 and DirectRT software on a computer equipped with a PST response box for button-presses and a microphone for voice-key.

**Results 2a: Self-paced Reading**

Lme models were fitted to reading times in Experiment 2 in a 6 (trial: 1-6) x 2 (context: informative v. neutral) design. Participants and items were included in each model as crossed random effects.

In total reading times on the target words, the analysis revealed a main effect of trial order such that for both informative and neutral contexts, reading times on novel words decreased over several exposures. The analysis revealed no effect of sentence contextual constraint and no interaction (see Table 10 for significance values and Table 11 for means).

Table 10

*Mixed-effects Models for Reading on the Target Word in Experiment 2.*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Fixed effect</th>
<th>b</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Reading Time</td>
<td>Trial</td>
<td>-103</td>
<td>5.12</td>
<td>-20.1</td>
<td>.0001</td>
<td>****</td>
</tr>
<tr>
<td>Target Word</td>
<td>Informative</td>
<td>-36.6</td>
<td>28.9</td>
<td>-1.27</td>
<td>.206</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>5.79</td>
<td>7.24</td>
<td>0.80</td>
<td>.424</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Significance codes: ‘†’ < .1, ‘*’ < .05, ‘**’ < .01, ‘***’ < .001, ‘****’ < .0001.
Table 11

*Mean Target Word Reading Times by Trial Order in Experiment 2.*

<table>
<thead>
<tr>
<th>Measure/Condition</th>
<th>Statistic</th>
<th>Trial Order</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total Time Target</td>
<td></td>
<td>Neutral</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>1353</td>
<td>1049</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>89</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Informative</td>
<td>1255</td>
<td>1096</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>69</td>
<td>64</td>
</tr>
</tbody>
</table>

Note. Reading times are given in milliseconds. SE = standard error of the mean.

**Results 2b: Post-reading Tests of Word Learning**

Logistic models were fitted to post-reading test accuracy measures and a cumulative link mixed model was fitted to the spelling recognition error data for the four nominal response types. As there was only a six exposures condition, the fixed effect in each model was for context type (neutral vs. informative). Subjects and items were included in each model as crossed random effects.

**Meaning recognition vocabulary test results.** Not surprisingly, the analysis for vocabulary test accuracy revealed a main effect of sentence contextual constraint ($b = 2.069, SE = 0.086, z = 24.16, p < .0001$) such that accuracy was higher for words seen in informative contexts than for words seen in neutral contexts. More interestingly, one-sample t-tests revealed that meanings for novel words seen six times in informative
contexts were recognized significantly above chance and that “meanings” for the novel words seen in neutral contexts were recognized at chance accuracy levels, as predicted (t(51) = 4.67, p < .0001, and t(51) = 0.22, p = .829, accordingly; see Table 12 for means). This finding replicates and extends Experiment 1. This is taken as evidence that the low accuracy rates in Experiment 1 were indeed due to making the task too difficult (i.e., there were too many new words to be learned in one sitting).

Table 12

Vocabulary Test Accuracy by Context Condition in Experiment 2.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Vocabulary Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>11.5 (0.3)</td>
</tr>
<tr>
<td>Informative</td>
<td>44.7 (0.4)</td>
</tr>
</tbody>
</table>

*Note.* Accuracy is in percentages. Standard errors are in parentheses. Chance levels of accuracy were 12.5%.

Because the meaning recognition task was too difficult in Experiment 1 and because there was not a better test of meaning recognition from a one exposure condition in Experiment 2, a test of retention for word meanings from one exposure using the new test method was needed. That is, can readers also demonstrate learning of meaning at above chance levels from one exposure? Or have previous studies using one exposure to words during reading shown an effect of semantic learning that is merely an artifact of the way the tests of memory in those studies were designed? In the test of memory for word meaning used in Experiments 1 and 2, readers would have had to tie form and
meaning together during reading to demonstrate learning of meaning after the reading session. Thus, in a separate experiment using pen-and-paper methods to provide this test of meaning from one exposure, eighty-nine Kent State University students participated for course credit. Participants were native English speakers and reported no reading disabilities. None had participated in any other portion of this study. Each participant read through 24 informative context sentences with 24 different novel words embedded (24 of the 32 novel words used in Experiment 1). Each of the novel words had different intended definitions. These 24 experimental sentences were intermixed with 24 filler sentences and the measure of interest was a two-choice version of this meaning recognition task. In the meaning recognition task, for example, both *blaph* and *flurst* had as meaning options *instrument* and *herb*. Again, in this way, differentiation was required and an episodic choice was disallowed. Overall, participants performed at better than chance (50%) accuracy on the meaning recognition test with 24 items ($M = 58\%, SE = 2\%$, $t(88) = 5.10, p < .0001$). Thus, readers demonstrate knowledge for word meaning at above chance levels from one exposure to novel words during reading. This evidences that readers do tie a meaning and spelling together during reading by inferring a meaning from context cues (without explicit definitions) and can demonstrate this after only one exposure.

**Spelling recognition test results.** A conditional analysis revealed no effect of sentence contextual constraint (or vocabulary accuracy in the informative condition) on spelling accuracy ($b = -0.081, SE = 0.073, z = -1.1, p = .27$; see Table 13 for means).
Table 13

*Spelling Accuracy and Target Word Reading Time as a function of Vocabulary Test Accuracy in Experiment 2.*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Spelling Accuracy</th>
<th>Reading Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral (no meaning available)</td>
<td>74 (3.3)</td>
<td>970 (54)</td>
</tr>
<tr>
<td>Informative (recognized meaning)</td>
<td>77 (4.1)</td>
<td>1050 (54)</td>
</tr>
<tr>
<td>Informative (did not recognize meaning)</td>
<td>68 (3.4)</td>
<td>876 (55)</td>
</tr>
</tbody>
</table>

*Note.* Accuracy is in percentages. Reading time is in milliseconds. Standard errors are in parentheses.

Thus, the prediction that novel word spellings read in informative contexts would be recognized with higher accuracy than spellings read in neutral contexts was not supported. In reading times on the novel word (Table 10) the model revealed no effect of contextual constraint on reading times for novel words ($p = .206$). However, as was suggested in Experiment 1, it is possible that reading times account for differences in response proportions on the spelling recognition test (i.e., longer reading times result in higher spelling recognition accuracies). An analysis of spelling recognition accuracy as a function of target word reading times supported this explanation ($b = 0.0002$, $SE = 0.00006$, $z = 3.37$, $p = .0008$; see Table 13 for means).

Thus, reading exposure impacted learning of word spelling, but learning of word meaning did not benefit learning of word spelling. However, did learning of word spelling benefit learning of word meaning? In investigating this, an analysis of accuracy on the Experiment 2 vocabulary test (only for words seen in informative contexts) as a
function of spelling recognition accuracy revealed that readers were significantly more accurate at identifying the meaning of a word when they had earlier correctly recognized its spelling ($b = 0.278$, $SE = 0.113$, $z = 2.45$, $p = .014$, see Table 14 for means). Further, reading times did not account for these differences ($b = -0.00005$, $SE = 0.0001$, $z = -0.29$, $p = .771$). Thus, learning of a new word’s orthography (and not reading times) benefitted learning of that word’s meaning.

Table 14

*Vocabulary Test Accuracy as a function of Accuracy on the Spelling Test in Experiment 2.*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Vocabulary Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral (no meaning available)</td>
<td>12 (1.7)</td>
</tr>
<tr>
<td>Informative (recognized spelling)</td>
<td>48 (3.8)</td>
</tr>
<tr>
<td>Informative (did not recognize spelling)</td>
<td>37 (5.5)</td>
</tr>
</tbody>
</table>

*Note.* Accuracy is in percentages. Standard errors are in parentheses. Chance accuracy on a test with 8 choices was 12.5%.

*Response comparisons.* In the cumulative link mixed model, there was again no effect of sentence contextual constraint on the frequency with which participants chose among response options ($b = 0.034$, $SE = 0.134$, $z = 0.26$, $p = .797$; see Table 15 for means), consistent with findings from Experiment 1. As in Experiment 1, from four response types, readers selected the correct spelling most often, followed by the homophone lure, followed by the transposition lure, followed by the transposed homophone. Among lures, participants were more likely to choose a Homophone Lure
than either of the transposed options \((t(51) = 10.2, p < .0001)\), with no difference in frequency of choosing between transposed options \((t < 1)\).

As is evident from Tables 7 and 15 (Six Exposures columns), reducing the number of new word spellings to be learned from 32 to 16 spellings increased spelling recognition accuracy across experiments.

Table 15

*Means for Spelling Recognition Test Responses in Experiment 2.*

<table>
<thead>
<tr>
<th>Response Type</th>
<th>Example Item</th>
<th>Six Exposures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Spelling</td>
<td>blaph</td>
<td>72.7 (2.3)</td>
</tr>
<tr>
<td>Homophone Lure</td>
<td>blaff</td>
<td>19.7 (1.5)</td>
</tr>
<tr>
<td>Transposition Lure</td>
<td>balph</td>
<td>4.9 (0.8)</td>
</tr>
<tr>
<td>Transposed Homophone</td>
<td>balff</td>
<td>2.6 (0.8)</td>
</tr>
</tbody>
</table>

*Note.* Means are in the percentage of the total responses belonging to each response type (adding up to 100%). Standard errors are in parentheses.

**Naming time task results.** Naming time data were again coded by removing errors and responses that were 2 standard deviations above or below each participant’s reaction time means. Data loss affected all conditions equally, 5.2% of responses were removed due to errors, and 3.4% of responses were removed due to falling outside 2 standard deviations of participant means.

In naming time data, words seen in informative contexts were expected to be named faster and with higher accuracy than words seen in neutral texts. In Experiment 2 naming times, the main effect of sentence contextual constraint did not approach
significance ($b = -9.053, SE = 6.840, t = -1.32, p = .186$). However, an analysis of pronunciation accuracy revealed a significant main effect of contextual constraint such that words seen in informative sentences were pronounced with fewer errors than words seen in neutral sentences ($b = 0.41, SE = 0.159, z = 2.59, p = .01$, see Table 16 for means). This is evidence that encountering words in informative contexts gives rise to a higher likelihood of pronouncing words correctly than encountering words in neutral contexts, though the effect is numerically very small.

Table 16

*Naming Times and Pronunciation Accuracy by Context Condition in Experiment 2.*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Naming Time</th>
<th>Pronunciation Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>753 (22)</td>
<td>93.9 (0.3)</td>
</tr>
<tr>
<td>Informative</td>
<td>739 (19)</td>
<td>95.7 (0.4)</td>
</tr>
</tbody>
</table>

*Note.* Naming times are in milliseconds. Accuracy is in percentages. Standard errors are in parentheses.

The main effect of sentence contextual constraint in pronunciation accuracy supported the prediction that word names that were tied to meaning during reading are given with higher accuracy than for words seen in neutral texts. However, conditional analyses for naming time data could again not be made as they were for spelling recognition data based on the very small number of cases in which a participant made an error in the naming time task in Experiment 2.

After removing words found to have the most variability in responses in Experiment 1 (those with the most errors), and providing the unused words in a practice
naming time session before the experimental naming time session, readers in the current experiment named each of the target words more quickly and accurately than they had in Experiment 1. Another way of testing the notion that phonological representations strengthen with exposure and become more stable over time (i.e., they would be named faster with more exposures) is to compare naming times for words read in Experiment 2 with those of a control group that only participated in the naming time task (after a practice naming session). Thus, a separate group of 13 Kent State University students\(^5\) participated for course credit in a control naming time study. None had participated in any other portions of this study (i.e., they had never seen the to-be-named words before). Participants were native English speakers and reported no reading disabilities. An independent samples \(t\)-test revealed that participants in Experiment 2 who read novel target words six times in sentences (experimental group participants) named those words significantly faster after the reading session than control group participants who did not read the words before the naming time task \((t\ (63) = -2.25, p = .044\), see Table 17 for means\). This finding provides further evidence that readers build phonological representations over several reading exposures (i.e., this representation of form builds stability with exposure).

\(^5\) The rule for data collection for the control group was to collect as many observations as possible before the end of the semester.
Table 17

*Table 17*

*Naming Times for Novel Words by Participants in Experiment 2 versus in a Control Group.*

<table>
<thead>
<tr>
<th>Group</th>
<th>Naming Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>745 (19)</td>
</tr>
<tr>
<td>Control</td>
<td>836 (49)</td>
</tr>
</tbody>
</table>

*Note.* Naming times are in milliseconds. Standard errors are in parentheses.

**Discussion**

In Experiment 2, the total reading time pattern on the target words replicated effects of trial order from Experiment 1 fixation time measures. Also, the reduction in the number of words to be learned in Experiment 2 was effective in allowing participants to demonstrate knowledge of word meanings after one reading session. The selection of the items for which the fewest pronunciation errors were made did reduce variability in naming time outcomes. Across two experiments, of predicted outcomes, only naming time differences were not observed (although small pronunciation accuracy differences were observed). Naming times were only faster for participants that saw the words during reading versus a control group of participants that did not see the words before the naming time task. A reason for this could be that with only six exposures in a single reading encounter, the orthographic strings are still relatively new. Thus, naming a new orthographic string seen once versus six times is relatively the same in terms of naming latencies. Future research will need to explore this over more exposures within a single session and over multiple sessions.
Experiment 2 provided evidence that tying form to meaning during reading benefits memory for both the form and meaning of novel words. Analyses revealed effects of sentence contextual constraint in phonological learning such that readers better learned new word pronunciations when those pronunciations were tied to meaning. However, learning of word spelling was not benefitted by learning of word meaning. As Webb (2008) suggested, it was reading exposure and not context quality for the sentences in which the new spellings were embedded during reading that impacted accuracy of orthographic knowledge after the reading session.

As in Experiment 1, error data from the spelling recognition task extended the pattern found by Yang and Perfetti (2006) such that readers more often chose a homophone lure than a transposed spelling containing the same letters as the correct spelling. This finding suggests that readers generate and develop phonological representations over several reading exposures and later access that phonological information in recognizing spellings of novel words. Again, another explanation for the reason that homophone lures were selected more often than transposed versions of the correct spellings could be that homophone lures were significantly more orthographically familiar than transposed lures. Thus, bigram frequencies were again compared between choice conditions to investigate this explanation in Experiment 2. Bigram frequencies did not differ between the homophone lures and transposition lures \( (t (15) = 1.70, p = .105) \) or between any of the four response options \( (\chi^2 = 4.88, p = .181; \text{see Table 18 for means}) \). Thus, means for token bigram frequencies across conditions in Experiment 2 further
suggest that readers use newly generated phonological representations for novel words in recognizing the spellings of those novel words.

Table 18

Means for Token Bigram Frequencies of Spelling Recognition Test Responses in Experiment 2.

<table>
<thead>
<tr>
<th>Response Type</th>
<th>Example Item</th>
<th>Bigram Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Spelling</td>
<td>glurch</td>
<td>464 (46)</td>
</tr>
<tr>
<td>Homophone Lure</td>
<td>glirch</td>
<td>544 (90)</td>
</tr>
<tr>
<td>Transposition Lure</td>
<td>grulch</td>
<td>379 (36)</td>
</tr>
<tr>
<td>Transposed Homophone</td>
<td>grilch</td>
<td>616 (84)</td>
</tr>
</tbody>
</table>

*Note.* Token bigram frequencies were calculated using the CELEX lexical database (Baayen et al., 1995). Standard errors are in parentheses.

Finally, results from Experiment 2 indicated that readers can actually learn the meanings for novel words from reading and demonstrate knowledge of semantics for these new words on a test of meaning recognition that was not confounded with episodic memory for task material. Specifically, vocabulary accuracy was significantly above chance levels following either one or six exposures to novel words from reading.
GENERAL DISCUSSION

Results from two experiments illustrated three major findings. First, post-reading measures indicated that learning of a novel word form benefits learning of that word’s meaning. Second, readers demonstrated effects of number of reading exposures and sentence contextual constraint in reading times for novel words. Third, on post-reading tests of learning for new word form and meaning, readers demonstrated above-chance levels of accuracy after only one exposure to novel words, and that accuracy increased after 6 exposures to novel words. These findings from reading times and post-reading tests of memory for the forms and meanings of new vocabulary words have practical, educational, and theoretical implications for word learning from reading.

Lexical Quality in Orthographic and Phonological Learning

As Balass et al. (2010) have shown with ERPs and Bolger et al. (2008) have shown in Experiment 1 (but not Experiment 2) of their study, it was expected that readers in the current study would demonstrate that learning of a new word meaning benefitted learning of word form. Pronunciation accuracies were higher for words seen in informative than neutral sentences, providing support for this prediction with phonology, but spelling accuracy was lower for words seen in informative than neutral contexts, not providing support for this prediction in orthographic learning (although this difference might be the result of a difference in intentional versus incidental word learning methodologies). Instead, and perhaps a more interesting and novel finding, readers demonstrated that learning of new word spelling benefitted learning of new word
meaning. Specifically, vocabulary accuracy was higher for words whose spellings were correctly recognized than words whose spellings were not correctly recognized after the reading session. This finding is consistent with the LQH (Perfetti & Hart, 2001), as this theory holds that high levels of individual word knowledge are the result of strong links between word meaning and word form.

In Experiment 1, readers demonstrated orthographic learning above chance from one exposure to novel words during reading and, after several exposures, readers demonstrated significant increases in their knowledge of word form. This finding using an incidental approach is consistent with evidence from direct instruction that for learning of word form, knowledge proceeds in small increments and LQ improves gradually over multiple exposures during reading (see Perfetti, 2007 for a review). This finding is also contrary to low orthographic accuracy rates using cued recall methods (e.g., Brusnighan et al., 2014), but is consistent with similar methods of spelling recognition that have been used previously (Bolger et al., 2008; Webb, 2008). Reading and spelling measures from Experiments 1 and 2 confirm that exposure (both the number of reading encounters and reading time spent during each encounter) impacted learning of word spelling, but context quality did not, as Webb (2008) concluded with participants learning English as a foreign language.

In the current study, both naming time and spelling recognition measures illustrated that readers use phonological information in learning new vocabulary words from silent reading. First, participants in Experiment 2, after encountering novel words during a reading session, named those novel words significantly faster after the reading
session than control group participants who did not read the words before the naming
time task. Second, spelling recognition data in both experiments mirrored that of Yang
and Perfetti (2006) such that foils on the spelling test were chosen most often for high
phonemic overlap, followed by high orthographic overlap, followed by foils with low
orthographic and low phonemic overlap. Together, this is taken as evidence that readers
create phonological representations for novel words during reading and later access these
phonological representations in demonstrating orthographic learning. This finding
extends previous fixation time and ERP evidence that phonological information is
activated during silent reading (Ashby, 2010; Ashby & Clifton, 2005; Brusnighan et al.,
2014; Folk, 1999; Folk & Morris, 1995). These findings are consistent with research
demonstrating that children’s phonological decoding skill predicts orthographic learning
(Bowey & Miller, 2007; Cunningham, 2006; Cunningham et al., 2002; Kyte & Johnson,
2006; Nation et al., 2007; Ricketts et al., 2011; Share, 1999). Share (1995) suggested that
establishing an orthographic representation for a novel word is dependent on
phonological analysis of that word in developing readers. As response comparisons in
Experiments 1 and 2 suggest, skilled adult readers’ orthographic learning may also be
dependent on phonological analysis.

**Effects of Contextual Constraint and Reading Exposure in Reading Novel Words**

Researchers have previously shown that fixation durations on a known word are
shorter on the second presentation of the word within a passage of text (Binder & Morris,
1995; Raney & Rayner, 1995; Rayner et al., 1994). In Experiment 1, readers
demonstrated effects of repetition that extended previous findings in two ways. Reading
times decreased with each of several exposures to novel words during reading in a single reading session. Readers also demonstrated effects of contextual constraint during reading such that total reading times for novel words in informative contexts were shorter than for novel words in neutral contexts. This extends reading patterns for known words preceded by constraining semantic contexts (see Clifton et al., 2007 for a review) to reading of novel words preceded by constraining contexts. That is, after one exposure in context, reading times for novel words in informative contexts were shorter than for novel words in neutral contexts because there is likely an emerging lexical representation for the novel words in informative contexts. Specifically, reading times were faster for novel words in informative contexts than for novel words in neutral contexts over several exposures because there is likely some preactivation of the emerging lexical representation with each new context consistent with that novel word’s meaning (after the first encounter in which a meaning was inferred).

**Post-reading Tests of Semantic Learning from Incidental Exposure during Silent Reading**

Fixation time measures have been used to demonstrate that readers can infer meanings for novel words from a single encounter with novel words in context during reading. Readers can subsequently demonstrate memory for these novel word meanings above chance accuracy levels from one reading encounter on post-reading vocabulary tests (Brusnighan & Folk, 2012; Brusnighan et al., 2014; Williams & Morris, 2004). However, previous post-reading vocabulary tests have not required differentiation between definitions for words seen during the reading task. It could be the case that when
such a test is given, a single reading exposure is not sufficient to demonstrate semantic knowledge for new word spellings. In the current study, a test that required differentiation between definitions for words seen during the reading session was administered. On this new vocabulary test, readers in Experiment 2 again demonstrated semantic knowledge from single exposures to new words in context above chance levels. Thus, this new vocabulary test requiring readers to match new word spellings with their meanings shows that it is not the case that a single reading exposure is insufficient to demonstrate semantic knowledge for new word spellings.

Next, following several reading exposures to novel words in context (as compared with performance on this new vocabulary test following one reading exposure in context), readers in the current study demonstrated significant deviations from chance levels of accuracy in learning of new word meanings on a new test of semantic learning. This finding illustrates that readers build on their knowledge of word meanings throughout a reading session over multiple exposures in a variety of contexts, increasing the stability of emerging lexical representations from low quality representations to high quality representations, consistent with the LQH (Perfetti & Hart, 2001). Of practical importance, using this new vocabulary test, readers were able to demonstrate learning of new word meanings for specific new word spellings.
APPLICATIONS AND FUTURE DIRECTIONS

In this study, new study methods for measuring orthographic, phonological, and semantic learning were tested. Results from two experiments suggested that these new test methods better quantify the amount of each type of learning that is possible from either one or several reading exposures than test methods such as cued recall and meaning recognition (from a seen vs. unseen during reading choice). Future research in this area could investigate the number of exposures necessary for robust learning of form and meaning (i.e., how many exposures are needed before accuracy is close to 100% for the spelling and/or meaning of a word?). Future research in this area could also explore the durability of these new word representations by testing learning of novel words from several reading exposures over several reading sessions (i.e., how much word knowledge is retained after several days?). Further, learning of new word form and meaning could be extended beyond exposures to new words in single sentences to more naturalistic reading materials (e.g., new words appearing throughout a larger passage of text).

These methods also have applications for intentional vocabulary learning methods and for vocabulary instruction methods for beginning readers and skilled, adult readers. That is, they provide a meaningful test of learning of word meaning and form from incidental exposure (such that spellings are tied to meaning) as they evidence that learning of word form and meaning are above chance from one exposure. In a review of studies evidencing different types of learning using eye movements from 2000 to 2012, Lai et al. (2013) suggest that only recently have researchers realized the educational implications of eye movements in learning environments. Their review shows that the
majority of studies sampled restrict their discussion of implications to information processing. Therefore, many educational researchers do not make good use of existing eye movement research.

Fisher and Frey (2014) argue that vocabulary is not an isolated skill, yet in too many cases, vocabulary instruction is isolated from other aspects of instruction in schools such as content area learning. This occurs in spite of researchers having long advocated for instructional approaches that capitalize on many forces. That is, a good instructional approach would be one that teaches structural, contextual, and morphemic analysis skills using a variety of channels (such as oral learning, teaching of word appreciation, and leveraging texts to facilitate discussion and interaction). Thus, instructional approaches should incorporate teaching of word consciousness (Graves & Watts-Taffe, 2008) to make students aware of the importance of learning word meaning in learning of word form and vice versa. This should make apparent to students that they have the tools to self-regulate their awareness of knowledge for different aspects of a new vocabulary word (e.g., the morphemes of the word, the word’s sound, word meaning, word spelling).

Stanovich (1986) demonstrated Matthew effects in both reading and vocabulary development, such that children with advanced vocabularies and reading skills continue to develop their reading skills and build their vocabularies, while their less advanced peers continue to struggle (i.e., the rich get richer and the poor get poorer). Kile (2013) suggests that teaching vocabulary acquisition strategies to beginning readers is the key to solving this problem. That is, educators should incorporate research into their instructional paradigms which emphasize that contextual, morphemic, phonemic, and
structural analysis skills can be combined to best benefit memory for all aspects of new vocabulary words from independent reading. In illustrating the importance of teaching vocabulary acquisition strategies to beginning readers, the current study methods should be adapted for beginning readers. In this way, previous findings demonstrating effects of sentence contexts in vocabulary learning from reading between second and sixth graders (West, Stanovich, Feeman, & Cunningham, 1983) can be extended.
REFERENCES


APPENDIX A

EXPERIMENT 1 SENTENCE FRAMES

6 EXPOSURE INFORMATIVE CONDITION:

(INSTRUMENT)
The musician played the _____ all morning.
The beautiful melodies came from the _____ in the hall.
The distant music was resonating from the _____ in the park.
The bridge of the song was played on a _____ last night.
Difficult notes can be learned with practice on the _____ over time.
The solo at the concert was performed on a _____ yesterday.

(GRAIN)
The farmer grows crops of _____ every year.
The fields are filled with rows of _____ in the summer.
The crows in the country can ruin a landscape of _____ unfortunately.
The harvester planted all of the _____ for a living.
The cereals and breads are made with ground _____ in them.
The farm equipment is made for harvesting _____ in the fall.

(INSECT)
The dog was bitten by a small _____ outside.
The exterminator sprayed to kill each _____ yesterday.
The walls were infested with a type of _____ that day.
The couple left out food that attracts the _____ to it often.
The fruit bat is considered useful each time it eats a _____ in cities.
The poison was formulated to get rid of the _____ from households.

(FRUIT)
The boy bit into the juicy _____ very slowly.
The bees were attracted to the ripe _____ that fell.
The tree drooped from the weight of each sweet _____ in the branches.
The mother had to peel each _____ for her children.
The baker made a pie filling with _____ yesterday.
The delicious smoothie contained chunks of real _____ for flavor.

(HERB)
The cook seasoned his sauce with _____ last night.
The secret ingredient in the soup is _____ at the restaurant.
The recipe calls for a hint of _____ at the end.
The spice rack included a shaker of _____ that was empty.
The best chefs flavor their dishes with _____ at home.
To flavor the stew he sprinkled _____ on it quickly.

(FISH)
The hawk circled the lake to catch and eat a _____ this morning.
The sportsman reeled in a huge _____ finally.
The captain said nets were great for catching a _____ out at sea.
The scales and bones should be removed before eating a _____ for dinner.
The aquarium is home to a beautiful _____ among favorites.
The oil spill at sea damaged the habitat of the _____ that lives there.
(TREE)

The boy sat under the shady branches of the _____ once a week.
The lumberjack struggled to chop down a _____ all morning.
The best firewood can be taken from the _____ to burn.
The cat had to be rescued from climbing a _____ this afternoon.
The most colorful leaves in the fall belong to the _____ this year.
The sticks in the yard all fell from the big _____ in the garden.

(PARTY)

Everyone wore a costume to the _____ last night.
The fraternity sent out invitations to the _____ this week.
The high school hosts at least one_____ yearly for students.
The DJ played music all night at the _____ for a fee.
The group wanted to join the festivities at the _____ last night.
The host had to ask guests to leave the _____ due to noise.
1 EXPOSURE INFORMATIVE CONDITION:

(STORE)
The patron gets whiskey from the _____ on Fridays.

(SMELL)
The perfume emits an enticing _____ in the store.

(KNIGHT)
The king was protected in battle by a faithful _____ each time.

(BELLHOP)
The salesman got a room key and gave his bags to the _____ to handle.

(CABINET)
The bachelor put cereal bowls away in the kitchen _____ this evening.

(DRINK)
The husband used a blender to mix his wife a _____ after dinner.

(FOOT)
The runner raced without shoes and hurt his left _____ in the race.

(STUNT)
The daredevil was famous for pulling off an amazing _____ last year.
6 EXPOSURE NEUTRAL CONDITION:

Every time Stephanie saw the _____ she was happy.

The people have heard about the _____ by now.

The newspaper will print a story about the _____ very soon.

There has been a _____ here for a long time.

Many people like the _____ already.

I think you will see the _____ eventually.

The man looked over at the _____ next to him.

A fairly common sight is the _____ around here.

Anna noticed there was a _____ in the distance.

A lot of people wondered about the _____ last month.

Sarah caught herself staring at the _____ momentarily.

Greg almost didn’t see the _____ because he wasn’t paying attention.

Sally was unhappy with the _____ yesterday.

You are likely to see the _____ sometime.

Brenda realized that there was a _____ right next to her.

Albrecht liked seeing the _____ earlier.

Marshall had been thinking about the _____ all day.

Shelly did not like encountering the _____ if she could help it.
Ally wanted to avoid dealing with the _____ at all costs.

Chris said that he had seen more than one _____ recently.

Amy couldn’t wait to see the _____ this summer.

Jerry had been excited about the _____ all year.

Brad wanted the _____ very badly.

Calvin was fond of the _____ that was there.

The girl looked at the _____ over her shoulder.

The picture of the _____ was fuzzy.

The painter captured the _____ perfectly.

Jessie thought that the _____ was awesome.

Bailey had never seen the _____ before today.

The boy paused at the sight of the _____ earlier.

Jan rarely saw the _____ until recently.

Tim thought about the _____ today.

The woman turned her head to look at the _____ today.

The writer mentioned the _____ in her book.

The light shined on the _____ today.

He went over to the _____ to take a picture.
She thought that the _____ was cool.

He casually glanced at the _____ near him.

She loved to see the _____ at times.

The picture showed the _____ entirely.

He focused on the _____ very closely.

She inspected the _____ that was in the painting.

I put my glasses on so I could see the _____ earlier.

He liked the _____ very much.

She watched the _____ for a while.

I didn’t like the _____ until today.

The girl examined the _____ yesterday.

He thought that the _____ was good.
EXPOSURE NEUTRAL CONDITION:

She was happy with the _____ earlier.

He hated the _____ quite a lot.

She never liked the _____ for some reason.

She saw the _____ out of the corner of her eye.

He had to deal with the _____ today.

Her friend always liked the _____ from the very beginning.

My friends are unsure how they feel about the _____ right now.

My uncle told me about the _____ yesterday.
## APPENDIX B

### EXPERIMENT 1 SPELLING RECOGNITION TEST

Circle the correct spelling for each novel word seen during the reading session.

<table>
<thead>
<tr>
<th>Choice Condition</th>
<th>Target Word</th>
<th>Homophone</th>
<th>Transposition</th>
<th>Transposed Homophone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>blaph</td>
<td>blaff</td>
<td>balph</td>
<td>baffle</td>
</tr>
<tr>
<td>2</td>
<td>flurst</td>
<td>flirst</td>
<td>frulst</td>
<td>frilst</td>
</tr>
<tr>
<td>3</td>
<td>ghoak</td>
<td>gohake</td>
<td>ghoake</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>gryle</td>
<td>grile</td>
<td>gyrle</td>
<td>girle</td>
</tr>
<tr>
<td>5</td>
<td>klaight</td>
<td>klate</td>
<td>kailght</td>
<td>kalte</td>
</tr>
<tr>
<td>6</td>
<td>kniest</td>
<td>kneast</td>
<td>kneist</td>
<td>knaest</td>
</tr>
<tr>
<td>7</td>
<td>chlorque</td>
<td>chlork</td>
<td>chrolque</td>
<td>chrolk</td>
</tr>
<tr>
<td>8</td>
<td>pleem</td>
<td>pleam</td>
<td>pelem</td>
<td>pelam</td>
</tr>
<tr>
<td>9</td>
<td>prooze</td>
<td>pruse</td>
<td>poorze</td>
<td>pusre</td>
</tr>
<tr>
<td>10</td>
<td>scraive</td>
<td>scrape</td>
<td>sacrive</td>
<td>sacrue</td>
</tr>
<tr>
<td>11</td>
<td>skeap</td>
<td>sleep</td>
<td>sekap</td>
<td>sekep</td>
</tr>
<tr>
<td>12</td>
<td>snoam</td>
<td>snome</td>
<td>sanom</td>
<td>somne</td>
</tr>
<tr>
<td>13</td>
<td>struice</td>
<td>stroose</td>
<td>scrutie</td>
<td>sortose</td>
</tr>
<tr>
<td>14</td>
<td>thraque</td>
<td>thrack</td>
<td>thorque</td>
<td>tharck</td>
</tr>
<tr>
<td>15</td>
<td>whyke</td>
<td>whike</td>
<td>whuke</td>
<td>whixe</td>
</tr>
<tr>
<td>16</td>
<td>wreize</td>
<td>wreeze</td>
<td>weirze</td>
<td>weirze</td>
</tr>
<tr>
<td>17</td>
<td>briece</td>
<td>breese</td>
<td>bierce</td>
<td>beerse</td>
</tr>
<tr>
<td>18</td>
<td>chright</td>
<td>chrite</td>
<td>chight</td>
<td>chirte</td>
</tr>
<tr>
<td>19</td>
<td>clurge</td>
<td>clirge</td>
<td>crulge</td>
<td>crilge</td>
</tr>
<tr>
<td>20</td>
<td>creeth</td>
<td>creath</td>
<td>ceerth</td>
<td>cearth</td>
</tr>
<tr>
<td>21</td>
<td>dwurt</td>
<td>dwert</td>
<td>druit</td>
<td>drewt</td>
</tr>
<tr>
<td>22</td>
<td>freague</td>
<td>frigue</td>
<td>feargue</td>
<td>firgue</td>
</tr>
<tr>
<td>23</td>
<td>glurch</td>
<td>glirch</td>
<td>grulch</td>
<td>grilch</td>
</tr>
<tr>
<td>24</td>
<td>gnize</td>
<td>gnuise</td>
<td>ginze</td>
<td>guinse</td>
</tr>
<tr>
<td>25</td>
<td>phrine</td>
<td>phreign</td>
<td>phrian</td>
<td>phriegn</td>
</tr>
<tr>
<td>26</td>
<td>quirth</td>
<td>querth</td>
<td>quirith</td>
<td>qureth</td>
</tr>
<tr>
<td>27</td>
<td>shoung</td>
<td>shongue</td>
<td>shoug</td>
<td>shogue</td>
</tr>
<tr>
<td>28</td>
<td>smaise</td>
<td>smaze</td>
<td>smiaze</td>
<td>smiaze</td>
</tr>
<tr>
<td>29</td>
<td>splorn</td>
<td>splorne</td>
<td>sproln</td>
<td>sproline</td>
</tr>
<tr>
<td>30</td>
<td>spruit</td>
<td>sprewt</td>
<td>spiurt</td>
<td>spwert</td>
</tr>
<tr>
<td>31</td>
<td>thweek</td>
<td>thwiek</td>
<td>theek</td>
<td>thiek</td>
</tr>
<tr>
<td>32</td>
<td>treave</td>
<td>treeve</td>
<td>tearve</td>
<td>teerve</td>
</tr>
</tbody>
</table>

*Note.* Trial order and choice order were randomized by the presentation software.
### APPENDIX C

#### EXPERIMENT 1 SPELLING RECOGNITION TEST SPELLING FREQUENCIES

<table>
<thead>
<tr>
<th>Target Frequency</th>
<th>Target Word</th>
<th>Spelling Body</th>
<th>Body Frequency</th>
<th>Homophone Lure</th>
<th>Spelling Body</th>
<th>Body Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>blaph</td>
<td>-aph</td>
<td>17</td>
<td>blaff</td>
<td>-aff</td>
<td>113</td>
</tr>
<tr>
<td>Low</td>
<td>flurst</td>
<td>-urst</td>
<td>33</td>
<td>flirst</td>
<td>-irst</td>
<td>1004</td>
</tr>
<tr>
<td>Low</td>
<td>ghoak</td>
<td>-oak</td>
<td>26</td>
<td>ghoke</td>
<td>-oke</td>
<td>273</td>
</tr>
<tr>
<td>Low</td>
<td>gryle</td>
<td>-yle</td>
<td>105</td>
<td>grile</td>
<td>-ile</td>
<td>919</td>
</tr>
<tr>
<td>Low</td>
<td>klaight</td>
<td>-aight</td>
<td>123</td>
<td>klate</td>
<td>-ate</td>
<td>1492</td>
</tr>
<tr>
<td>Low</td>
<td>kniest</td>
<td>-iest</td>
<td>16</td>
<td>kneast</td>
<td>-east</td>
<td>539</td>
</tr>
<tr>
<td>Low</td>
<td>chlorque</td>
<td>-orque</td>
<td>5</td>
<td>chlork</td>
<td>-ork</td>
<td>33</td>
</tr>
<tr>
<td>Low</td>
<td>pleem</td>
<td>-eem</td>
<td>230</td>
<td>pleam</td>
<td>-eam</td>
<td>285</td>
</tr>
<tr>
<td>Low</td>
<td>prooze</td>
<td>-ooze</td>
<td>6</td>
<td>pruse</td>
<td>-use</td>
<td>598</td>
</tr>
<tr>
<td>Low</td>
<td>scraive</td>
<td>-aive</td>
<td>1</td>
<td>scraive</td>
<td>-ave</td>
<td>507</td>
</tr>
<tr>
<td>Low</td>
<td>skeap</td>
<td>-eap</td>
<td>56</td>
<td>skeep</td>
<td>-eep</td>
<td>537</td>
</tr>
<tr>
<td>Low</td>
<td>snoam</td>
<td>-oam</td>
<td>43</td>
<td>snoame</td>
<td>-ome</td>
<td>569</td>
</tr>
<tr>
<td>Low</td>
<td>struice</td>
<td>-uice</td>
<td>13</td>
<td>strooese</td>
<td>-oose</td>
<td>60</td>
</tr>
<tr>
<td>Low</td>
<td>thraque</td>
<td>-aque</td>
<td>2</td>
<td>thrack</td>
<td>-ack</td>
<td>1525</td>
</tr>
<tr>
<td>Low</td>
<td>whyke</td>
<td>-yke</td>
<td>1</td>
<td>whike</td>
<td>-ike</td>
<td>1188</td>
</tr>
<tr>
<td>Low</td>
<td>wreize</td>
<td>-eize</td>
<td>6</td>
<td>wreeze</td>
<td>-eese</td>
<td>33</td>
</tr>
<tr>
<td>High</td>
<td>briece</td>
<td>-iece</td>
<td>141</td>
<td>breeese</td>
<td>-eese</td>
<td>3</td>
</tr>
<tr>
<td>High</td>
<td>chright</td>
<td>-ight</td>
<td>2529</td>
<td>chrite</td>
<td>-ite</td>
<td>904</td>
</tr>
<tr>
<td>High</td>
<td>clurge</td>
<td>-urge</td>
<td>33</td>
<td>clirge</td>
<td>-irge</td>
<td>2</td>
</tr>
<tr>
<td>High</td>
<td>creeth</td>
<td>-eeth</td>
<td>103</td>
<td>crest</td>
<td>-eath</td>
<td>12</td>
</tr>
<tr>
<td>High</td>
<td>dwurt</td>
<td>-urt</td>
<td>71</td>
<td>dwert</td>
<td>-ert</td>
<td>2</td>
</tr>
<tr>
<td>High</td>
<td>freague</td>
<td>-eague</td>
<td>69</td>
<td>frigue</td>
<td>-igue</td>
<td>1</td>
</tr>
<tr>
<td>High</td>
<td>glurch</td>
<td>-urch</td>
<td>351</td>
<td>glirch</td>
<td>-irch</td>
<td>2</td>
</tr>
<tr>
<td>High</td>
<td>gnize</td>
<td>-ize</td>
<td>169</td>
<td>gnruise</td>
<td>-uisse</td>
<td>6</td>
</tr>
<tr>
<td>High</td>
<td>phrain</td>
<td>-ain</td>
<td>690</td>
<td>phreign</td>
<td>-eign</td>
<td>7</td>
</tr>
<tr>
<td>High</td>
<td>quirh</td>
<td>-irth</td>
<td>76</td>
<td>quirth</td>
<td>-erth</td>
<td>4</td>
</tr>
<tr>
<td>High</td>
<td>shoung</td>
<td>-oung</td>
<td>385</td>
<td>shongue</td>
<td>-ounge</td>
<td>35</td>
</tr>
<tr>
<td>High</td>
<td>smaise</td>
<td>-aise</td>
<td>70</td>
<td>smaze</td>
<td>-aze</td>
<td>48</td>
</tr>
<tr>
<td>High</td>
<td>splorn</td>
<td>-orn</td>
<td>238</td>
<td>splorne</td>
<td>-orne</td>
<td>9</td>
</tr>
<tr>
<td>High</td>
<td>spruit</td>
<td>-uit</td>
<td>83</td>
<td>sprewt</td>
<td>-ewt</td>
<td>8</td>
</tr>
<tr>
<td>High</td>
<td>thweek</td>
<td>-eek</td>
<td>477</td>
<td>thwiek</td>
<td>-iek</td>
<td>5</td>
</tr>
<tr>
<td>High</td>
<td>treave</td>
<td>-eave</td>
<td>211</td>
<td>treeve</td>
<td>-eeve</td>
<td>11</td>
</tr>
</tbody>
</table>

*Note.* Spelling frequencies were taken from the Ziegler, Stone, and Jacobs (1997) database for computing feedforward and feedback consistency in English.
APPENDIX D

EXPERIMENT 1 VOCABULARY TEST EXAMPLES (conditions in parentheses for clarity, trials and choices were randomized during testing by the presentation software)
Select the correct meaning for each novel word seen during the reading session.

1. blaph (informative):
   a. instrument
   b. grain
   c. insect
   d. fruit
   e. herb
   f. fish
   g. tree
   h. party
   i. store
   j. smell
   k. knight
   l. bellhop
   m. cabinet
   n. drink
   o. foot
   p. stunt

2. creeth (informative):
   a. instrument
   b. grain
   c. insect
   d. fruit
   e. herb
   f. fish
   g. tree
   h. party
   i. store
   j. smell
   k. knight
   l. bellhop
   m. cabinet
   n. drink
   o. foot
   p. stunt
### flurst (neutral):

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>bird</td>
</tr>
<tr>
<td>b.</td>
<td>lawyer</td>
</tr>
<tr>
<td>c.</td>
<td>pancake</td>
</tr>
<tr>
<td>d.</td>
<td>airplane</td>
</tr>
<tr>
<td>e.</td>
<td>mushroom</td>
</tr>
<tr>
<td>f.</td>
<td>groundhog</td>
</tr>
<tr>
<td>g.</td>
<td>bag</td>
</tr>
<tr>
<td>h.</td>
<td>hammer</td>
</tr>
<tr>
<td>i.</td>
<td>chair</td>
</tr>
<tr>
<td>j.</td>
<td>table</td>
</tr>
<tr>
<td>k.</td>
<td>flower</td>
</tr>
<tr>
<td>l.</td>
<td>movie</td>
</tr>
<tr>
<td>m.</td>
<td>twig</td>
</tr>
<tr>
<td>n.</td>
<td>miner</td>
</tr>
<tr>
<td>o.</td>
<td>nurse</td>
</tr>
<tr>
<td>p.</td>
<td>neighbor</td>
</tr>
</tbody>
</table>

### glurch (neutral):

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>bird</td>
</tr>
<tr>
<td>b.</td>
<td>lawyer</td>
</tr>
<tr>
<td>c.</td>
<td>pancake</td>
</tr>
<tr>
<td>d.</td>
<td>airplane</td>
</tr>
<tr>
<td>e.</td>
<td>mushroom</td>
</tr>
<tr>
<td>f.</td>
<td>groundhog</td>
</tr>
<tr>
<td>g.</td>
<td>bag</td>
</tr>
<tr>
<td>h.</td>
<td>hammer</td>
</tr>
<tr>
<td>i.</td>
<td>chair</td>
</tr>
<tr>
<td>j.</td>
<td>table</td>
</tr>
<tr>
<td>k.</td>
<td>flower</td>
</tr>
<tr>
<td>l.</td>
<td>movie</td>
</tr>
<tr>
<td>m.</td>
<td>twig</td>
</tr>
<tr>
<td>n.</td>
<td>miner</td>
</tr>
<tr>
<td>o.</td>
<td>nurse</td>
</tr>
<tr>
<td>p.</td>
<td>neighbor</td>
</tr>
</tbody>
</table>
## APPENDIX E

### EXPERIMENT 2 SPELLING RECOGNITION TEST

Select the correct spelling for each novel word seen during the reading session.

<table>
<thead>
<tr>
<th>Choice Condition</th>
<th>Target Word</th>
<th>Homophone Lure</th>
<th>Transposition Lure</th>
<th>Transposed Homophone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>flurst</td>
<td>flirst</td>
<td>frulst</td>
<td>frilst</td>
</tr>
<tr>
<td>2</td>
<td>ghoak</td>
<td>ghoke</td>
<td>gohak</td>
<td>gohake</td>
</tr>
<tr>
<td>3</td>
<td>pleem</td>
<td>pleam</td>
<td>pelem</td>
<td>pelam</td>
</tr>
<tr>
<td>4</td>
<td>prooze</td>
<td>pruse</td>
<td>poorze</td>
<td>pusre</td>
</tr>
<tr>
<td>5</td>
<td>skeap</td>
<td>skeep</td>
<td>sekap</td>
<td>sekep</td>
</tr>
<tr>
<td>6</td>
<td>snoam</td>
<td>nome</td>
<td>sanom</td>
<td>somne</td>
</tr>
<tr>
<td>7</td>
<td>struice</td>
<td>stroose</td>
<td>scrutie</td>
<td>sortose</td>
</tr>
<tr>
<td>8</td>
<td>wreize</td>
<td>wreeze</td>
<td>weirze</td>
<td>weerze</td>
</tr>
<tr>
<td>9</td>
<td>briece</td>
<td>breese</td>
<td>bierce</td>
<td>beerse</td>
</tr>
<tr>
<td>10</td>
<td>creeth</td>
<td>creath</td>
<td>ceerth</td>
<td>cearth</td>
</tr>
<tr>
<td>11</td>
<td>dwurt</td>
<td>dwert</td>
<td>druitt</td>
<td>drewt</td>
</tr>
<tr>
<td>12</td>
<td>glurch</td>
<td>glirch</td>
<td>grulch</td>
<td>grilch</td>
</tr>
<tr>
<td>13</td>
<td>phrain</td>
<td>phreign</td>
<td>phrian</td>
<td>phriegn</td>
</tr>
<tr>
<td>14</td>
<td>quirth</td>
<td>querth</td>
<td>quirith</td>
<td>qureth</td>
</tr>
<tr>
<td>15</td>
<td>splorn</td>
<td>splorne</td>
<td>sproln</td>
<td>sprolne</td>
</tr>
<tr>
<td>16</td>
<td>treave</td>
<td>treeve</td>
<td>tearve</td>
<td>teerve</td>
</tr>
</tbody>
</table>

*Note.* Trial order and choice order were randomized by the presentation software.
## APPENDIX F

**EXPERIMENT 2 VOCABULARY TEST EXAMPLES** (conditions in parentheses for clarity, trials and choices were randomized during testing by the presentation software)

Select the correct meaning for each novel word seen during the reading session.

1. **briec** (informative):
   1. instrument
   2. grain
   3. insect
   4. fruit
   5. herb
   6. fish
   7. tree
   8. party

2. **creeth** (informative):
   1. instrument
   2. grain
   3. insect
   4. fruit
   5. herb
   6. fish
   7. tree
   8. party

3. **flurst** (neutral):
   1. soap
   2. lawyer
   3. hammer
   4. exercise
   5. building
   6. chair
   7. arm
   8. wrestler

4. **glurch** (neutral):
   1. soap
   2. lawyer
   3. hammer
   4. exercise
   5. building
   6. chair
   7. arm
   8. wrestler