Copyright 2012 Kacey L. Wochna
COMPARING INCIDENTAL LEARNING OF NOUNS AND VERBS USING EYE MOVEMENTS

BY

KACEY L. WOCHNA

THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Educational Psychology in the Graduate College of the University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Master’s Committee:

Associate Professor Kiel Christianson, Chair
Professor Elizabeth A. L. Stine-Morrow
Associate Professor Kara D. Federmeier
Abstract

It has long been assumed that people learn much of their vocabulary incidentally during the course of natural reading. Prior research has identified a host of factors that affect the ease or likelihood of learning a new word from written context. Findings from various areas of cognitive science suggest that part of speech should be one such factor. Nouns and verbs differ in their syntactic and semantic roles. As a result of these differences, verbs are generally considered to be harder to process and harder to learn. Most incidental word learning studies are not appropriate for examining part of speech differences, and those that have tried have found inconclusive results. In the present study, the effect of part of speech on incidental word learning was isolated and examined. Participants read real and novel nouns and verbs in one and two single sentence contexts while their eyes were tracked. Post-tests of word knowledge indicated that nouns were learned more effectively than verbs. The reading times for sentences containing novel nouns and verbs reflected the differences between these word classes. Readers devoted more attention to contextual information when a novel word was a verb. Reading patterns were also distinctly affected by multiple exposures to the novel word in either the same context or different contexts.
# Table of Contents

**Chapter I: Introduction and Literature Review** .......................................................... 1  
  Learning from Context ................................................................. 3  
  Influences on Contextual Word Learning ................................................. 8  
  Part of Speech ............................................................................. 11  
  Present Study ............................................................................. 14  

**Chapter II: Experiment 1** ............................................................................. 18  
  Method ....................................................................................... 18  
  Results ....................................................................................... 24  
  Discussion ................................................................................... 29  

**Chapter III: Experiment 2** ............................................................................. 32  
  Method ....................................................................................... 33  
  Results ....................................................................................... 36  
  Discussion ................................................................................... 40  

**Chapter IV: General Discussion** ..................................................................... 44  
  Limitations and Future Directions ...................................................... 45  
  Conclusion ................................................................................... 48  

References .......................................................................................... 50  

Tables and Figures ........................................................................... 57  

Appendix ............................................................................................ 65
Chapter I

Introduction and Literature Review

A large vocabulary is commonly touted as one of the benefits of reading. For some time, researchers also assumed that this relationship was true because, as Jenkins, Stein, and Wysocki (1984) point out, there was no other explanation for the amount of vocabulary growth seen in school-age children. Spoken language does not contain enough rare words to be a useful source of vocabulary after early childhood – the average conversation between college graduates contains fewer rare words than the average children’s book (Cunningham & Stanovich, 2001).

Most schools devote very little time to explicit vocabulary instruction and when they do, it is not particularly effective (Beck, McKeown, McCaslin, & Burkes, 1979; Scott, Jamieson-Noel, & Asselin, 2003). Even the most rigorous programs cover only a few hundred words a year (Nagy & Herman, 1984). Students are not adept at using and learning from dictionaries, so it is unlikely that much vocabulary is learned from looking up the meanings of unfamiliar words (Scott & Nagy, 1997). People also continue to learn new words into adulthood. The idea that people learn new words unintentionally during the course of normal reading became the default explanation because “the other contenders have dropped from the race” (Jenkins et al., 1984, p. 769).

Early research about the phenomenon of incidental word learning first had to establish that it was possible for people to learn new words from natural reading. Nagy and colleagues conducted multiple studies showing that incidental word learning does occur at a modest but still potentially powerful rate. Given the rate of word learning found in their study, Nagy, Herman, and Anderson (1985) calculated that incidental word learning could account for the amount of vocabulary growth seen in children, which has been estimated as at least 3,000 words a year. Although it would be difficult to show definitively that this is how vocabulary is learned,
researchers have felt comfortable concluding that incidental word learning probably is the source of much of adults’ vocabulary knowledge. Subsequent research has focused on what the process of gradually learning a word’s meaning looks like, and the factors that affect how well or quickly words are learned.

One factor that may influence the odds of incidentally learning a word is part of speech. The syntactic and semantic differences between nouns and verbs have been well researched throughout the fields of psychology and linguistics, but have been largely ignored as a source of variation in incidental learning. The materials in incidental word learning experiments typically come from pre-existing texts; while this is good for ecological validity, it is often not possible to assess word-level effects such as part of speech independently because these effects are naturally confounded. The purpose of the present study was to isolate and investigate the effect of part of speech on incidental word learning using a psycholinguistic approach.

In the present study, adults read novel nouns and verbs, as well as familiar nouns and verbs, in informative single sentence contexts while their eye movements were recorded. They were then tested on the meanings of the novel words. If verbs are more difficult to learn from context than nouns, as previous research from other fields would predict, performance on the tests would be expected to be better for nouns than verbs. The reading time information was not used to measures learning, but to examine how novel words affect the reading process in comparison with real words. If part of speech affects how easily a word is learned from context, novel nouns and verbs would also be expected to evoke different reading time patterns.

This chapter provides relevant background literature for this study. First, I present a general overview of the nature of incidental word learning. This is followed by a review of the characteristics of learners, contexts, and words that have been shown to influence how well
words are learned. Finally, I discuss research on part of speech from various areas of cognitive science and the implications of this research for incidental word learning.

**Learning from Context**

It is not always obvious when a study concerns incidental word learning. Many relevant studies use different language to describe the same process. Conversely, some reports discuss their findings as though they reflect incidental learning when their experiments arguably measure something different. In their meta-analysis of incidental word learning research, Swanborn and de Glopper (1999) define incidental word learning as “incidental, as opposed to intentional, derivation and learning of new word meanings by subjects reading under reading circumstances that are familiar to them” (p. 262). In order to learn a new word incidentally, the reader must be able to infer something about the meaning of that word from its context. Context is usually understood as the written text surrounding a word, from the sentence containing the unknown word on a local level, to the entire document (a paragraph, a passage, a book, etc.) on a global level. However, because incidental word learning assumes a normal reading situation, the mental representation of a text that is constructed as part of the comprehension process can also be considered context (Diakidoy, 1998).

Swanborn and de Glopper (1999) emphasize that *incidental* is in contrast with *intentional* because many studies have focused on the ability of readers to intentionally derive the meanings of words in contexts. Some of these studies have not made a clear distinction between this activity and incidental word learning. In studies of intentional word learning, participants are instructed to either (a) derive the meanings of unknown words and provide definitions while the supporting context is still in front of them, or (b) use contextual information to actively try to
learn unknown words, which requires remembering the meanings for later testing. Although incidental word learning requires some derivation of meaning, readers in a derivation task may use strategies that are not triggered by a natural reading situation (Nagy, Anderson, & Herman, 1987). For example, Freebody and Anderson (1983) found that it took an incredibly high proportion of rare words in a passage (1 in 3) before comprehension was negatively affected, and hypothesized that either natural prose is so redundant that guesses about meaning are very accurate even in the presence of unknown words, or that readers may skip unknown words (and maybe even the propositions containing them) entirely in order to commit minimal effort to comprehension.

Speculations that novel words in context receive little attention have not been supported by subsequent research. Most explicitly, eye-tracking studies have indicated that readers notice unfamiliar words in text and exert more effort in processing them (Chaffin, Morris, & Seely, 2001; Williams & Morris, 2004). It is still assumed, however, that readers will exert more effort deriving the meanings of unknown words if they are instructed to do so than if left to their own devices. Konopak et al. (1987) attempted to compare incidental and intentional word learning directly: Half of the participants read a passage where the unfamiliar words were underlined while the other half read the same passage with no underlining. Readers of the underlined passage did better when they were later asked to define those words than readers of the plain passage, although readers of the plain passage exhibited some learning. However, the participants in the intentional learning group also completed an additional definition task prior to the post-test. It is possible that it was this extra practice, and not intentionality alone, that improved performance.
In their meta-analysis, Swanborn and de Glopper (1999) were particularly interested in determining the probability of learning a new word from context. They determined that the average probability across the 20 studies they included in their analysis was 15%, but many studies have observed much lower rates than this. For example, Nagy et al. (1987) calculated the probability of learning a new word from context to be 5%. Many factors can influence that amount of word learning that occurs during a study. One that Swanborn and de Glopper point out is whether or not a study gives credit for partial word knowledge. Incidental word learning, and in fact word learning in general, is not all or nothing. Even when readers are intentionally deriving the meaning of an unknown word, it may be impossible to determine the precise meaning of that word from the context alone. Schatz and Baldwin (1986) examined the plausibility of incidental word learning and determined that naturally occurring prose is not very helpful as a means of deriving an unknown word’s meaning.

The assumption that is made in claiming that people learn a large portion of their vocabulary from context is that they may need to read a word many times in a variety of different contexts before they learn its meaning. Behind this assumption is the assumption that word learning of any kind is rarely all or nothing, but instead is incremental. Meaning representations are shaped over multiple encounters with a word (Frishkoff, Collins-Thomas, Perfetti, & Callan, 2008; Frishkoff, Perfetti, & Collins-Thompson, 2011). Many traditional tests of vocabulary require a reader to have a fairly complete representation of a word. Some researchers who are specifically interested in partial word knowledge have used a vocabulary checklist developed by Shore and Durso (1990) called the Level of Word Knowledge Assessment Task (LOWKAT). In addition to known and unknown words, their test includes the classification of frontier words. A frontier word is only partially known: The reader acknowledges familiarity with the word, but
otherwise cannot define it or use it appropriately in context. For the LOWKAT, participants are given a list of words and nonwords and complete instructions requiring decreasing levels of word knowledge. Words that are correctly defined or used in a sentence are considered known. Words marked as familiar are considered frontier. Real words marked as unreal words are considered unknown. Instruments that can assess levels of partial knowledge make it possible to examine whether the development of word meaning representations is systematic. In one example of this, Whitmore, Shore, and Hull Smith (2004) compared knowledge about taxonomic associates and thematic associates. They found that while people had some knowledge of both at all levels, taxonomic information was already highly available at low knowledge levels and changed less with increasing knowledge than thematic knowledge.

There is a need for even more sensitive measures. Durso and Shore (1991) found that even when participants could not identify that a word existed in the English language (the criterion for being unknown), they could accurately perform tasks which required some knowledge of meaning. They could distinguish between correct and incorrect use of the unknown word in a sentence as long as the uses differed in general (e.g., animacy vs. inanimacy) rather than specific constraints (e.g., for the target happy: *The happy child played in the park* vs. *She put the vase on the happy table*); between a word semantically related to the unknown word and a semantically unrelated word (e.g., *glad* vs. *purple*); and between a word that could sensibly appear with the unknown word and a word that could not (e.g., *happy child* vs. *happy table*). On some tasks, participants performed as well with unknown words as they did with the familiar frontier words. Durso and Shore also tested whether knowledge about unknown words was based on the words’ orthography/phonology or on concepts that could not have been derived from the
words’ physical characteristics: Performance levels were equally high regardless of whether word knowledge was physically or conceptually based.

Complete meaning representations typically require many encounters with a new word in a variety of contexts, but studies have shown that it is possible to acquire a measurable degree of word knowledge from a single exposure to an unknown word (Chaffin et al., 2001; Diakidoy, 1998; Nagy et al., 1987; Nagy et al., 1985; Williams & Morris, 2004). Learning from a single exposure has been demonstrated under a variety of conditions. Diakidoy and Nagy and colleagues presented unknown real words in naturally occurring, passage-long contexts, often textbook or story excerpts. While this is most typical, Chaffin et al. (2001) presented unknown words (nonwords replacing known words) in informative, two-sentence-long contexts and Williams and Morris (2004) presented unknown nonwords in informative single sentence contexts. These studies also differ in how they measured learning. On the lenient side of the scale, participants in Williams and Morris’s study only had to discriminate between two possible categories for each learned word, while on the stricter side, participants in Nagy et al. (1987) had to pick the correct definitions from three semantically similar distractors.

Studies vary tremendously in the materials they use and the tasks they present, and not all researchers have successfully found learning after a single exposure. Jenkins et al. (1984) gave participants two, six, or ten contexts to read for each of the unknown words in their experiment, and then measured word knowledge with three tests of varying difficulty. They found that it took six exposures before participants performed above chance levels on any of the tests, and only after ten exposures were they able to demonstrate word knowledge at all three levels. Wagovich and Newhoff (2004) had participants read two stories containing eight unknown words each. Each unknown word occurred only once in the text. The participants then answered multiple
choice questions about the words’ definitions that varied in difficulty: Some required only knowledge of the word’s part of speech, some required only general category information, and some required detailed semantic knowledge. The results of the multiple-choice post-tests revealed no learning at any difficulty level. The only measured benefit in this single exposure experiment was that after reading the words in context, readers labeled more of them correctly as real English words. These findings suggest that the likelihood of acquiring a measurable degree of knowledge about a new word from a single exposure is highly dependent on the materials, the learning task, and the assessment.

Influences on Contextual Word Learning

The study of incidental word learning has been complicated by the variability introduced into the learning situation by characteristics of the people doing the learning, the contexts that contain them, and the unknown words themselves.

Educators are particularly interested in the individual differences that affect the ability to learn words from context, specifically reading ability. Because vocabulary is thought to affect comprehension, there is a concern that good comprehenders also learn more new words from context, resulting in a Matthew effect (e.g., Jenkins et al., 1984). Not all studies have found that reading ability predicts learning (Nagy et al., 1987; Nagy et al., 1985; Wagovich & Newhoff, 2004), but general opinion holds that readers with higher verbal ability (whether that is assessed by reading skill, vocabulary knowledge, or some other measure) are better at learning new words from context than readers who are less skilled (Herman, Anderson, Pearson, & Nagy, 1987; Jenkins et al., 1984; Swanborn & de Glopper, 1999). This is not surprising, given that reports linking vocabulary and comprehension can be found within the incidental word learning
literature in addition to an extensive body of work outside the scope of this review (Diakidoy, 1998; Jenkins et al., 1984).

Daneman and Green (1986) and McGinnis and Zelinski (2000) both found that working memory capacity, as measured by reading span (e.g., Daneman & Carpenter, 1980) was correlated with ability to derive word meanings from context. This result most likely extrapolates to incidental learning, and may account for part of the predictive effect of verbal ability. How exactly verbal ability and working memory interact with learning from context is an open question.

Many researchers have assumed that word learning is largely dependent on the nature of the context, especially the extent to which the context is informative about the meaning of an unknown word. Contexts are generally considered highly informative if they contain explicit clues to a word’s meaning, such as synonyms or short definition, if they present a word in a way that strongly limits, or constrains, the number of plausible meanings, or if they generally provide a lot of information about a word’s meaning. This is a reasonable intuition: In order to learn a new word from context alone, it must be possible to infer something about its meaning from the context. Borovsky, Kutas, and Elman (2010) found that brain responses reflected learning of novel words when they had been read in informative single-sentence contexts but not when they had been read in uninformative contexts. However, most manipulations of contextual informativeness have had little effect on learning (Diakidoy, 1998; Jenkins et al., 1984; Nagy et al., 1987; Schwanenflugel, Stahl, & Mc Falls, 1997). This finding suggests that incidental learning may be less like intentional learning than expected, as contextual informativeness, conceptualized as close, explicit meaning clues and highly constraining phrases, unambiguously supports intentional learning (e.g., Daneman & Green, 1986; Frishkoff et al., 2008).
When word learning is a by-product of comprehending, it may be advantageous to consider the mental representation of the text (i.e., the situation model) to be the informative context. Studies of passage-level conceptual difficulty paint a picture of word learning as highly dependent on a learner’s pre-existing conceptual network (Diakidoy, 1998; Herman et al., 1987; Nagy et al., 1987). A new word can be learned while reading if the learner already has the background knowledge necessary to support its meaning. If the text concerns unfamiliar concepts, however, the task of learning new content and new vocabulary may be too overwhelming. Swanborn and de Glopper (1999) likewise found that less was learned from texts that had high densities of unknown words in their meta-analysis. In their study, Nagy et al. found that when passage-level conceptual difficulty was controlled, informativeness became a significant predictor. They suggested that normally, more informative texts might also be more conceptually difficult such that the benefit of the former is negated by the effect of the latter.

Independent of learner ability and context quality, some words are just more difficult to learn. Concreteness is a quality of words that reflects how easily the concepts they represent can be experienced by the senses: Words for objects or features that are physical and can be seen are very concrete (such as table) while words that cannot be conventionally sensed are not concrete or abstract (such as justice). Abstract words are harder to learn than concrete words (Schwanenflugel et al., 1997). Nagy et al. (1987) found that conceptually simple words are easier to learn than conceptually difficult words. By their definition, conceptual difficulty depends on the prior conceptual knowledge of an individual, although generalizations can be drawn across people based on what is likely to be common knowledge for a population. The most conceptually simple word represents a known concept that has a one-word synonym (altercation). A more conceptually difficult word represents an unknown concept that can be understood based on
known concepts (naïve). The most conceptually difficult word represents an unknown concept that requires learning new information or a related set of concepts (ventricle). Some word factors that might be expected to effect learning, such as length and morphological transparency, have not been found to have an effect (Nagy et al., 1987). The current study focuses on one prominent word characteristic whose effect on learning has yet to be determined: part of speech.

**Part of Speech**

Findings from research on spoken language acquisition, psycholinguistics, and cognitive neurophysiology all strongly predict that nouns should be easier to learn incidentally than verbs. Young children’s vocabularies contain a disproportionately large number of nouns in terms of both comprehension and production (e.g., Gentner, 1978; Goldin-Meadow, Seligman, & Gelman, 1976). In addition to being learned later, children continue to make semantic errors when using very common verbs long after many infrequent nouns have been mastered (Bowerman, 1982). In a study where new words were explicitly taught, Childers and Tomasello (2002) found that 2-year-olds learned more nouns than verbs. Additionally, in simulations of early learning, adults and older children are much better able to identify nouns than verbs from watching silent videos of mothers interacting with their children (Gillette, Gleitman, Gleitman, & Lederer, 1999, Piccin & Waxman, 2007). In Gillette et al., the effect of part of speech could be completely explained by differences in imageability, confirming the assumption that if a learner must depend solely on observations of the physical world, words that correspond to physical things will be easier to learn. The first words a child learns must be learned through extralinguistic observation, and it is easier to map nouns to their referents in the world than to map verbs to their relational components (Gleitman, 1990).
This difference in the ease of making word-to-world mappings should not directly affect learning from written context, because there is no physical world to interpret. However, there is further evidence that differences between nouns and verbs persist long after language is acquired. Genter (1981) suggests that the ease and consistency with which words can be mapped to the physical world actually cause nouns (particularly concrete nouns) to be fundamentally different from verbs in nature. She posits that the internal perceptual elements of object nouns are closely bound, while verbs are defined less by internal components than by links to external objects. In other words, the perceptual world tightly constrains how objects are lexicalized, and less tightly constrains how relational information is lexicalized.

Gertner’s (1981) proposal is evidenced by cross-linguistic comparisons. It is well known that languages differ in which meaning components they include in their verbs. Gentner offers English and Spanish motion verbs as an example: In English, motion verbs incorporate the manner of motion but relegate directional information to adverbs or prepositions (e.g., run into the house), while Spanish encodes directional information in the main verb but not manner (e.g., entrar en la casa corriendo, literally enter the house running) (Talmy, 1978, cited in Genter, 1981).

This difference explains Genter’s (1981) finding that when phrases that are translated from one language into another and then back again by two different bilingual speakers, the verbatim nouns are almost always preserved, but the verbs are likely to be changed in the process. She further argues that the somewhat unstable nature of verb meaning helps explain why memory for verbs is poorer than for nouns, why verbs have a wider breadth of meaning (common verbs have substantially more senses than common nouns), why the frequency distribution of verbs is top-heavy (more like closed-class words than like nouns), and why people
are more liberal in interpreting verbs than nouns in semantically mismatched sentences (Gentner & France, 1988). Given their comparatively nebulous nature, it would not be surprising if verbs were harder to learn incidentally than the generally more semantically “fixed” nouns.

The most explicit evidence that verbs are probably harder than nouns to learn incidentally comes from a comparison of nouns and verbs using event-related brain potentials (ERPs; Federmeier, Segal, Lombrozo, & Kutas, 2000). They found that people had greater N400 responses to pronounceable pseudowords in contexts when they were used as verbs than when they were used as nouns. The N400 is an ERP component that is associated with semantic processing, and it is generally larger for words that are unpredictable or semantically unrelated to previous words. Federmeier et al. thus concluded that the semantic processing of novel verbs is more difficult than novel nouns.

Few incidental word learning studies have explicitly examined whether part of speech affects learning. Nagy et al. (1987) included part of speech as one of a number of possible word-level predictors of learning, but found no effect. Schwanenflugel et al. (1997) found that children were worse at incidentally learning nouns than verbs or adjectives. They hypothesized that nouns may be difficult for literate children to learn because of conflicts with heuristics originally used when acquiring spoken language. For example, young children tend to assume that new labels refer to new whole objects, not to parts or features of objects, but as they learn less common words, more of these words will in fact refer to parts or aspects, requiring an adjustment in strategy. The findings of Shore and Kempe (1999) suggest that adults have different heuristics that they use for interpreting unfamiliar words. They examined how people are able to choose correct contexts for words whose meanings they do not know. They found that people could use semantic domain size, the number of words that exist to describe a concept, to help make these
decisions. People assume that unknown words come from large semantic domains: For example, *clamor* is more likely to mean *noise* than *cheese grater* because there are many words that also mean *noise*. It seems plausible that with extensive experience with the language, people transition from one set of heuristics to another. In that case, however, Schwanenflugel et al.’s finding would not be expected to apply to adult learners. The Scwanenflugel et al. report also found the previously stated effect of concreteness, but they did not control for the concreteness or imageability of the nouns and verbs in their study. Although the authors do not say, the nouns were probably more concrete on average, which makes the finding of a verb advantage even more surprising.

Wagovich and Newhoff (2004) found that after reading unknown nouns and verbs in context, children correctly labeled more verbs as real words, but showed no improvement on nouns. Because this result was unexpected, they mention that some of the verbs had prefixes, which may have acted as memory cues, and that all of the verbs were marked infinitives but only half of the nouns were preceded by a determiner. They speculated that readers might perceive verbs as more important for comprehension and thus devote more effort to them, but any additional effort must have been slight because performance on two tests of meaning knowledge showed no learning for either syntactic category. Together, these studies provide only scant evidence that verbs are easier to learn from written context.

**Present Study**

The experiments in the present study directly compare the incidental learning of nouns and verbs. On average, nouns and verb in natural prose differ in concreteness, which is known to affect learning, and probably differ in the typical informativeness or constraint of their contexts,
which may result in a bias in readers’ incidentally acquired vocabularies. The purpose of the present study was to isolate the effect of part of speech on learning and answer the following question: If concreteness and contextual informativeness were equal, would one still be more difficult to learn than the other?

The basic design of these experiments was modeled after the method used by Williams and Morris (2004). In their study, participants read a collection of sentence contexts containing either low-frequency words or novel pseudowords while their eye movements were tracked. Eye-tracking has not commonly been used to study incidental learning, but it can provide valuable information about how unknown words affect reading on-line. Eye movement measures may reveal differences in how noun and verbs (and their contexts) are processed, even in the absence of any measurable learning. After reading the novel words in their contexts, participants were tested on the meanings of the novel words. Williams and Morris found that readers spent more time reading the novel words and their contexts than reading low-frequency words and their contexts. They also found that novel words whose meanings were correctly selected on the post-test were read more quickly initially but re-read for more time than words whose meanings were not correctly selected.

Like in Williams and Morris (2004), participants in the present experiments read real words and novel pseudowords in single-sentence contexts while their eye movements were tracked. The contexts were structured similarly, with neutral context preceding the real and novel targets and highly informative context following them. However, all of the targets used by Williams and Morris were nouns: In these experiments, half of the targets were verbs.

The post-test used by Williams and Morris (2004) required participants to pick the correct category for each novel word from two choices. Although participants in their study saw
each novel word only once, and in a minimal amount of context, performance on this post-test was better than chance. Given that their method resulted in measurable learning, I opted to employ a similar test for semantic meaning. Keeping in mind that the ability to learn new words from a single exposure seems to be highly situational, a second post-test that only required remembering the part of speech of the novel words was used so that there would be another avenue for assessing learning. Because part of speech was also the factor of interest, differences between nouns and verb might be apparent even at this level of word knowledge.

The contexts used by Williams and Morris (2004) were sufficiently informative that readers could derive and recall the category membership of the novel words. It is unclear what level of specificity readers will be able to learn about verbs in similar contexts. The verb materials in the present experiments were selected to approximately match the noun materials in hierarchical position (i.e., superordinate, basic, subordinate) but while the taxonomic structure for nouns is generally agreed upon (e.g., Malt, 1995; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), the equivalent system for verbs is less familiar. The lack of obvious hierarchy may be one more factor that contributes to the relative difficulty of learning verbs.

In addition to being similarly specific and taxonomically located, the nouns and verbs were also all highly imageable. To the extent that it was possible control, their contexts were equally informative. By controlling these factors, I increased certainty that any differences would be due to word class. Because a noun advantage was so strongly predicted, only transitive verbs were used to try to maximize the amount of semantic and syntactic information available to learners. Even with this accommodation, I expected nouns to be learned better than verbs. The reading time patterns were more difficult to predict. If unknown verbs are harder to learn, it is reasonable to expect that readers will spend more time trying to process them than unknown
nouns. On the other hand, the relational nature of verbs may result in increased effort reading the context but not the novel words themselves.
Chapter II

Experiment 1

Experiment 1 used the method developed by Williams and Morris (2004) to investigate whether part of speech affects the ease of incidentally learning a new word from context. Participants read highly informative sentences containing real nouns, real verbs, novel nouns, and novel verbs while their eye movements were recorded. The real nouns and verbs were matched on a number of characteristics, including length and frequency. The materials were constructed so that it across participants, it was possible to compare eye movement measures on sentences that differed only in whether the target word was real or novel. It was also possible to compare eye movements for the same novel word form used as a noun and as a verb. I expected that novel words would be read for longer than their real counterparts, as they were in Williams and Morris. Comparisons of noun and verb reading patterns were more exploratory. After the reading task, participants were tested on their memory for the part of speech and semantic category of each novel word. I predicted that performance on both tests would be better for nouns than verbs.

Method

Participants. Thirty-three members of the University of Illinois community participated in the experiment. Participants were compensated with course credit or $7. All participants were native English speakers and had normal or corrected-to-normal vision.

Materials. Twenty nouns and twenty verbs were used as the real word targets. Words were selected using WordNet (Princeton University, 2010), a lexical database of super-subordinate relations. For nouns, ten categories of approximately equal specificity were selected.
Two hyponyms were then selected from each category. A similar process was conducted for the verbs, using the hierarchical relationship of troponymy: While a hyponym noun is a more specific type of another noun (a daisy is a type of flower), a troponym verb is a specific manner of another verb (to smack is to hit in a specific manner). Ten verb categories were selected, with two troponyms per category. All verbs were transitive. Targets were selected so that each noun matched a verb exactly in number of characters (range 4 – 9, $M = 5.40$), and to within 10 occurrences per million words in frequency (range 1 – 52; nouns: $M = 15.60$; verbs: $M = 14.35$; Davies, 2008).

It was not possible to match words on imageability or concreteness, both because norms were not available for all words, particularly verbs, and because the nouns are generally rated as more imageable and concrete. Matching them would have required selecting nouns with lower imageability and concreteness, and this was too difficult within the category and frequency constraints. Instead, verbs were only selected if I judged them to be highly imageable. The average imageability and concreteness ratings for the verbs (based on 7 ratings) were 528.7 and 490.1, while the average rating for the nouns (based on 16 ratings) were 594.6 and 604.9. Both ratings range from 100 to 700, and were collected from the MRC Psycholinguistic Database (Coltheart 1981), which merges three sets of norms (Gilhooly & Logie, 1980; Pavio, Yuille, & Madigan, 1968; Toglia, & Battig, 1978).

Each noun-verb pair was then matched with a pronounceable pseudoword of the same character length. These pseudowords served as the novel targets. Although they were orthographical regular, with an average type bigram frequency of 23.15 (compared with 34.59 for nouns and 41.61 for verbs), most had no orthographic neighbors ($M = .35$ neighbors; bigram
frequencies and neighborhood density collected from N-Watch; Davis, 2005). A list of all the target words is provided in Table 1.

The noun and verb targets were each embedded into a unique single sentence frame, 76 to 80 characters in length. The target was never less than three words from the beginning of the sentence and never more than five. These contexts were neutral preceding the targets, and informative for the meaning of the target after the target. The predictability of the preceding context and the informativeness of the proceeding context was normed in a separate experiment described below.

**Norming.** The sentence contexts were normed by 35 participants who did not participate in either of the main experiments. Participants were provided with 74 sentence frames up to the target and were asked to fill in the next word, to ensure that the targets were not predictable. These participants were also provided with the complete sentence frames with the targets removed and were asked to fill in the blank, to ensure that the meanings of the targets could reliably be determined from context. Frames were immediately discarded if more than half of the participants selected a word in the same category as the target in the predictability task, if no participants selected a word in the same category as the target in the informativeness task, or if all participants selected the exact target word in the informativeness task. The remaining items were selected such that the noun and verb frames matched as closely as possible in informativeness. The average predictability of the final 40 contexts was .03 for the nouns and .05 for the verbs, which was not a significant difference ($t = 0.95$, $df = 36$, $p = .346$). The average informativeness was .93 for the nouns and .85 for the verbs, a marginally significant difference ($t = -1.83$, $df = 37$, $p = .075$). A list of all experimental items is included in the Appendix.
**Design.** Figure 1 presents the design used in Experiment 1. The sets of real nouns and their frames, real verbs and their frames, and novel words were separated into two groups (A and B) such that each group contained only one member from each noun or verb category (e.g., *daisy* was in group A, so *tulip* was in group B; see Table 1 for a list of targets in each group). Whether a frame was seen with its real target or its corresponding novel target was counterbalanced across two lists. For example, a participant in list 1 saw (1) the real nouns from group A in their frames; (2) the real verbs from group B in their frames; (3) the novel words from group A in the group A verb frames; and (4) the novel words from group B in the group B noun frames:

1. Jason saw a daisy growing near the mailbox and he picked it for his mother.
2. Tim tried to punch William during their fight but his fist met the wall instead.
3. The teachers sometimes pavif their students with rulers when they misbehave.
4. Claire wished she had a betik to add to the bouquet but all of hers were wilted.

Each participant thus saw ten frames containing real nouns, ten containing real verbs, ten containing novel nouns, and ten containing novel verbs for a total of 40 experimental items per participant. As indicated in the examples above, each participant saw only one novel word for each noun or verb category; the other category member was seen in its real word form. Across all participants, all frames were seen with both real and novel targets, resulting in a 2 (novelty: real or novel) x 2 (part of speech: noun or verb) final design.

**Post-tests.** Two post-tests were created to evaluate whether participants learned the novel targets. For the part of speech test, each novel target was presented in two content-empty sentence frames, used as a noun in one and as a verb in the other, to test whether participants
learned and remembered syntactic information about the novel targets. The part of speech post-test choices for the novel word *pavif* are presented in (1). If the participant saw *pavif* in a noun frame, sentence (1a) would be the correct response. If they saw *pavif* in a verb frame, sentence (1b) would be the correct response.

(1) a) I will have a pavif.
    b) He will pavif him.

For the category test, each novel target was presented with two possible meanings, the correct category for that target and one of the other categories from the same part of speech, to test whether participants remembered semantic information about the novel targets. Category post-test choices for the novel word *pavif* are presented in (2). When seen as a noun, *pavif* was in the frame for *daisy*, so category (1b) would be the correct response.

(2) a) vegetable
    b) flower

The same part of speech post-test was used for all participants because the choices were appropriate regardless of list: The correct answer for one list was the incorrect answer for the other list. Because the category choices were dependent on whether the novel word was seen in a noun or verb frame, separate category post-tests were made for each list. For example, participants in list 2 saw *pavif* as the verb *smack*, so the choices on their test were *hit* (correct) and *cook* (incorrect).
Apparatus. Eye movements were recorded by a SR Research Ltd. Eyelink 1000/2000 eye tracker, which records the position of the reader’s eye once every millisecond (1000 Hz sampling rate), and has a high spatial resolution of 0.01°. Participants were seated approximately 72.5 cm away from a 20-inch monitor where the sentences and paragraphs were displayed. Stimuli were displayed in 14 pt. Courier New font. At this distance, 3 characters subtended 1 degree of visual angle. A chin rest and forehead rest were used to reduce head movements. Although viewing was binocular, eye movements were recorded from only the right eye.

Procedure. Participants were run one at a time, and the experiment lasted approximately 30 minutes. At the beginning of the eye-tracking session, a 9-point calibration was performed. Participants were instructed to fixate nine circles that appeared around the screen. If the calibration was deemed inaccurate, a new calibration routine was conducted. Calibration accuracy was monitored during the course of the experiment by the experimenter, and a new calibration was performed whenever accuracy was deemed inadequate.

Participants were given verbal instructions to read each sentence silently and for comprehension. No mention was made of vocabulary learning or novel words, nor were participants told that they would be tested after the eye-tracking portion. Prior to each sentence, participants had a one-point calibration check. If the calibration was deemed accurate, a sentence appeared on the computer screen. When finished reading, participants pressed a button on a control pad, causing the sentence to disappear and be replaced by another one-point calibration check. Each participant read 40 experimental sentences and 40 filler sentences in pseudorandomized order. Yes-or-no comprehension questions were presented visually after 82.5% of the fillers to ensure that they remained on task. Participants responded by pressing one
of two buttons on the control pad. They did not receive feedback on their accuracy. The average accuracy on these questions was 94%.

After completing the eye-tracking portion, the post-tests were administered. First, participants rated the familiarity of the novel words. Participants were given a list including all of the novel targets and ten additional pseudowords that were not read during the eye-tracking portion and asked to mark whether they had seen the word before, including during the experiment. Participants were then given the part of speech test. They were instructed to select which of two sentences was using the novel target correctly, a task that required only knowledge of the novel target’s part of speech. Finally, they were given the category test. Participants were instructed to select the word that best described the novel targets from two possible category choices.

Results

Eye movements were analyzed for four different sentence regions: the target region, the pre-target region, the post-test region, and the sentence-ending region. The focal region was the target region, which consisted of just the target word. The pre-target region consisted of the word immediately preceding the target word. The post-target region varied depending on part of speech. For nouns, it consisted of the word immediately following the target word. For verbs, it began with the word immediately following the target and ended with the direct object of the verb phrase. It ranged from one to five words in length, but most often consisted of just an article and a noun. This was done in order to better examine the use of the relational structure in the verb sentence, for which there was no obvious correlate in the noun sentences. The sentence-ending region consisted of all words following the post-target region.
The pre-target, post-target, and sentence-ending regions of the noun and verb sentence frames differed in number of characters, making any comparison based on untransformed reading time measures would have been difficult to interpret. One solution to this problem, proposed by Trueswell, Tanenhaus, and Garnsey (1994), is to subtract the actual reading times from the predicted reading times for that string length for each participant. These residual reading times were calculated by fitting a separate linear mixed model for each reading time measure, with length (in characters) of the region as the fixed effect and with random intercepts and slopes for each participant. Residual reading times were used as the dependent variable in the models in place of the raw reading times, making direct comparison of the noun and verb sentences feasible.

Up to seven eye movement measures were examined for each region. *First fixation duration* is the duration of the first fixation made in the target region. *Gaze duration*, or first-run time, is the summed duration of all fixations in a region from when it is first entered until it is first exited. *Go-past time* is the summed duration of all fixations in a region from when it is first entered until it is first exited to the right. *Total time* is the summed duration of all fixations in a region during the trial. *Regressions in* reflects whether a regression was made into the region from a later part of the sentence. *Regressions out* reflects whether a regression was made out of the region to an early part of the sentence. *Skipping* reflects whether the region was fixated. The four duration variables are in milliseconds, while the other three are binomial (1 if a regression was made or a region was skipped, 0 if otherwise).

Duration measures were analyzed using linear mixed effect models. Models were fit in the R software package (R Development Core Team, 2011) using the lmer() function of the lme4 package (Bates, Maechler & Bolker, 2011). All models included novelty, part of speech, and
their interaction as fixed effects, and random intercepts for participants and items. The lmer() function does not provide p-values for effects in linear models. With data sets of this size, however, it is appropriate to consider fixed effects to be significant if the absolute value of their t-statistic is greater than two (Baayen, Davidson, & Bates, 2008). The binomial measures were similarly analyzed using logit mixed models, with novelty, part of speech, and their interaction as fixed effects, and random intercepts for participants and items. Effect significance was determined using the p-values provided by the analyses.

Accuracy and familiarity rating were binomial measures. For accuracy, 1 represents an accurate response, and 0 represents an inaccurate response. Because all test questions had two answer choices, 50% accuracy indicates chance performance. For familiarity ratings, 1 represents a response indicated that a word looked familiar, and 0 represents a response indicating that a word looked unfamiliar. All three were analyzed using logit mixed models. For part of speech accuracy and category accuracy, part of speech, accuracy on the other post-test, and familiarity rating were included as fixed effects, with random intercepts for participants and items. For familiarity rating, the model included part of speech and accuracy on both tests as fixed effects, and random intercepts for participants and items.

**Target region.** Meaning reading times on the target region are provided in Tables 2 and 3. All seven eye movement measures were analyzed for the target region. Table 4 displays the estimates, standard errors, and t-values for the model parameters. There was an effect of novelty for all seven measures: When the target was novel, reading times were longer, regressions in and out were more likely, and less skipping occurred. The effect of part of speech was never a significant predictor. The interaction between novelty and part of speech had an effect for all measures except regressions out. For first fixation duration, gaze duration, and go-past time,
when the target was real, verbs had longer reading times than nouns, but when the target was novel, nouns had longer reading times than verbs. Similarly, real verbs were more likely to be regressed into and less likely to be skipped than real nouns, while novel nouns were more likely to be regressed into and less likely to be skipped than novel verbs. The pattern was slightly different for total time: Real verbs had longer total times than real nouns, but when the target was novel, total time on nouns and verbs did not differ.

**Pretarget region.** Total time and regressions in were analyzed for the pre-target region. Total time was effected by part of speech \((b = 48.90, SE = 24.35, t = 2.01)\) and novelty \((b = -39.40, SE = 17.20, t = -2.29)\). Pre-target had longer reading time before verb targets than noun targets, and before novel targets than real targets. There was an effect of novelty on regressions in \((b = -.66, SE = .20, z = -3.26, p = .001)\), but the effect of part of speech was marginal \((b = .37, SE = -.20, z = 1.86, p = .06)\).

**Post-target region.** Go-past time, total time, and regressions out were analyzed for the post-target region. Novelty had an effect on go-past time \((b = -64.45, SE = 14.41, t = -4.47)\), total time \((b = -83.81, SE = 19.55, t = -4.29)\), and regressions out \((b = -1.37, SE = .26, z = -5.19, p < .001)\). The post-target region had longer reading times and was more likely to be regressed out of when the target was novel rather than real. There was also an effect of part of speech on go-past time \((b = 77.24, SE = 31.10, t = 2.48)\), total time \((b = 77.24, SE = 31.10, t = 2.48)\), and regressions out \((b = .54, SE = .22, z = 2.51, p = .012)\). Novelty and part of speech interacted in all three measures: go-past time \((b = -74.98, SE = 20.38, t = -3.68)\); total time \((b = -116.28, SE = 27.68, t = -4.20)\); and regressions out \((b = .70, SE = .32, z = 2.15, p = .03)\). For the duration measures, the pattern of this interaction was the reverse of the pattern on the target word. When the target was real, reading times on the post-target region were longer for nouns than verbs.
When the target was novel, post-target reading times were longer for verbs. For regressions out, the interaction reflects that the effect of novelty on the post-target was larger for noun targets than for verb targets, although verb targets were more likely to have regressions out of the post-target regardless of novelty.

**Sentence-ending region.** Total time and regressions out were analyzed for the sentence-ending region. In the sentence-ending region, there was an effect of novelty on total time ($b = -125.58$, $SE = 48.93$, $t = -2.57$), where the region had longer reading times when the target was novel. There was also an interaction effect ($b = -191.48$, $SE = 69.17$, $t = -2.77$): Like the post-target region, noun target sentences had longer reading times when the target was real, while verb target sentences had longer times when the target was novel. Sentences endings with novel targets were more likely to be regressed from ($b = -.69$, $SE = .18$, $z = -3.74$, $p = .0002$).

**Post-tests.** On the part of speech post-test, the accuracy for nouns was better than chance ($t = 2.33$, $df = 329$, $p = .02$; $M = .56$), but the accuracy for verbs did not differ from chance ($t = -1.44$, $df = 327$, $p = .15$; $M = .46$). Part of speech had an effect on accuracy ($b = -.41$, $SE = .17$, $z = -2.41$, $p = .02$) such that nouns had better accuracy than verbs. Familiarity rating also had an effect on accuracy ($b = .45$, $SE = .16$, $z = 2.83$, $p = .005$). Items rated as familiar were more likely to be answered accurately than those rated as unfamiliar.

On the category post-test, the accuracy for verbs was better than chance ($t = 2.11$, $df = 327$, $p = .04$; $M = .56$), but the accuracy for nouns did not differ from chance ($t = 0$, $df = 329$, $p = 1$; $M = .50$). However, the effect of part of speech was not significant ($b = .26$, $SE = .25$, $z = 1.06$, $p = .29$), and familiarity rating had only a marginal effect on accuracy ($b = .31$, $SE = .17$, $z = 1.82$, $p = .07$), so the above-chance performance on the verbs may not be reliable.
Nouns and verbs did not differ in their familiarity ratings \( (b = -.23, SE = .22, z = -1.05, p = .29; \) noun \( M = .45, \) verb \( M = .41) \). Length of the target had an effect on rating \( (b = -.38, SE = .09, z = -4.27, p < .0001) \) such that the shorter the word, the more likely it was to be rated as familiar.

Mean-centered gaze durations and total times for the various sentence regions were tested as possible predictors for the two test accuracies and familiarity ratings, but none of the eye movement measures improved these models and were thus were not included in the final fits.

**Discussion**

Experiment 1 found that more effort was exerted to read novel targets and their contexts receive than the same contexts containing real words. Given that previous eye-tracking studies have shown the same effect for low-frequency words (Chaffin et al., 2001; Williams & Morris, 2004), it is unsurprising that the higher-frequency words used in the present study induced the same effect. Far more interesting are the patterns of interactions between novelty and part of speech. When the target words were real, readers spent more time processing verbs than nouns. When the targets were novel, readers spent more time on nouns than verbs. This pattern appeared in measures of both initial reading (first fixation duration, gaze duration, skipping) and rereading (go-past time, regressions in).

This finding is not intuitive, but is more easily understood if compared with the reading times on the context regions. When targets were real, the post-target region and the sentence ending were read for more time when they contained noun targets than verb targets. When the targets were novel, the reverse was true: Contexts of verbs received more attention. Regressions out were always more likely out of post-targets following verbs, but the increase in regressing
out prompted by novelty was smaller than that for nouns. The pre-target word was reread more before verbs, regardless of novelty. The whole picture is one where readers encountering an unknown word focus their attention on the novel word itself when it is a noun, but on contextual information when the word is a verb.

These results seem consistent with Gentner’s description of verb semantics (1978; 1981; Gentner & France, 1988). In their proposal of the verb mutability hypothesis, Gentner and France emphasize that the conceptualization of verbs as less semantically stable is not necessarily in conflict with theories of verb centrality (e.g., Kintsch, 1974). Verb centrality theories claim that the verb is essential to language processing because it provides the relational framework for the phrase. The relative importance of the verb for comprehension may be the cause of longer reading times on real verbs. Gentner and France suggest that the importance of the verb’s linking function might actually be the cause of its mutability: When the semantics of nouns and verb conflict, providing relational structure takes precedence over maintaining the normal components of the verb’s meaning. It is plausible that a similar process happens when the verb’s semantics are absent, as in the case of novel words: Effort is first focused on establishing the relations between the other parts of the phrase, and hence increased reading of the context.

The idea that establishing the relational framework is the first step in verb learning is supported by research in other areas. There is substantial evidence that children do this when acquiring spoken language (Fisher, 1996; Fisher, 2002; Fisher, Gleitman, & Gleitman, 1991). The theory of syntactic bootstrapping claims that children use abstract knowledge of syntactic rules to learn new verbs. For example, children younger than three can use the number of noun phrase arguments in a sentence to determine the correct referent for a verb (Fisher, 2002). Although the adults in the study did not have a visual referent, meaning could be derived from a
written sentence in a similar way. According to Pustejovsky’s (1995) generative lexicon theory, a verb’s semantic representation is composed of an argument structure, which describes the number and type of the verb’s arguments, an event structure, which describes what kind of event the verb is, and qualia, which specify the aspects of the verb’s meaning. When readers encounter an unknown verb, the most immediately available component is its argument structure. Once a learner identifies the number and type of the verb’s arguments, the event structure may be discernible. Because the qualias of the noun phrase arguments specify how they relate to other words, it may be possible to make deductions about the verb’s qualia. If this is the process for learning verbs from written context, it is clear that the arguments, and hence the context, must be fully processed before the verb can be interpreted. The fact that novel nouns and verbs had the same total reading times provides further evidence that while learners devote extra processing to novel nouns when they first come to them, extra processing is only devoted to novel verbs themselves after their contexts have been examined.

The eye movement data from Experiment 1 shows a difference in how novel nouns and verbs are read. Unfortunately, it provided less information about whether those processing differences also lead to learning differences. Readers performed better for nouns than verbs on the part of speech test, but in general, accuracy for both parts of speech on both tests was very close to chance. Although William and Morris (2004) demonstrated that it is possible to measure knowledge of words learned in this way, a single exposure was not enough for learners in this experiment. Experiment 2 was designed to further facilitate learning so that part of speech differences could be detected.
Chapter III

Experiment 2

The purpose of Experiment 2 was to generate information about noun and verb learning to complement the information about noun and verb processing revealed in the first experiment. Because I hoped to boost performance on the post-tests while changing the method and materials as little as possible, I opted for the simplest alteration: having participants read each novel target twice. However, it seemed likely that merely having participants read the same sentences two times might not improve memory substantially – readers might decide to skip over sentences that they remembered having already seen. Even if participants reread the sentence, the exposure might boost recognition of the target without actually improving knowledge of its meaning because the second sentence would not provide any new information about the novel word. Additionally, repeating the same sentence verbatim would be quite unlike natural incidental learning. Normally, incremental learning results from seeing unknown words in a variety of different contexts. In order to improve post-test performance while testing these theories about incremental learning, two repetition conditions were employed in this experiment. In the same-frame repetition condition, participants read sentences containing novel words two times each. In the different-frames repetition condition, participants saw each novel word twice, but in different sentence frames. Because the materials from Experiment 1 contained two sentences frames for each noun and verb category, it was possible to create both of these conditions with only minimal changes to the materials.

Measures of individual differences were also added to Experiment 2. Although the overall accuracy in Experiment 1 was at chance, it is possible that some of the readers had learned the new words. Because verbal ability and working memory capacity are correlated with
incidental learning ability (e.g., Daneman & Green, 1986; Swanborn & de Glopper, 1999), collecting this information about participants would allow me to distinguish readers who are more and less likely to be successful learners. This would be particularly useful if learning in the overall sample is still too low to detect part of speech differences.

Method

Participants. Forty-three members of the University of Illinois community participated in the experiment. Participants were compensated with course credit or $7. All participants were native English speakers and had normal or corrected-to-normal vision. None had participated in Experiment 1 or the materials norming study.

Materials. The novel targets and sentence frames were the same as in Experiment 1. The real targets were not used.

Design. Figure 2 presents the designs used in Experiment 2. The novel words and frames were grouped as they were in Experiment 1. Instead of 40 target words, participants saw only the 20 novel words, but each novel word was seen twice. Participants were assigned to one of two repetition conditions. In the same-frame repetition, each novel word was used in one frame, and participants read that frame two times. Whether a novel word was seen in a noun frame or a verb frame was counterbalanced across two lists. For example, a participant in list 1 would see (1) the novel words from group A in the group A noun frames two times and (2) the novel words from group B in the group verb frames two times:

(1) Jason saw a pavif growing near the mailbox and he picked it for his mother.

Jason saw a pavif growing near the mailbox and he picked it for his mother.
(2) Tim tried to ruzel William during their fight but his fist met the wall instead.

Tim tried to ruzel William during their fight but his fist met the wall instead.

Each participant thus saw ten frames containing novel nouns and ten frames containing novel verbs. Because each frame was repeated, each participant still saw 40 experimental items but only 20 were unique. As in Experiment 1, each participant saw only one novel word for each noun or verb category: If a participant saw the frame for *daisy*, they did not see the frame for *tulip*. Across all participants in this condition, all novel words were read in both their noun and verb frames.

In the different-frames condition, each novel word was used in two frames that were seen once each. One frame was from the same group as the novel words, as has always been the case previously (e.g., *pavif* in the *daisy* frame). The other was the frame for the word of the same category in the opposite group (e.g., *pavif* in the *tulip* frame). Whether a novel word was seen as in a noun frame or a verb frame was counterbalanced across two lists. For example, a participant in list 1 would see (1) the novel words from group A in the group A noun frames; (2) the novel words from group A in the group B noun frames; (3) the novel words from group B in the group B verb frames; and (4) the novel words from group B in the group A verb frames:

(1) Jason saw a pavif growing near the mailbox and he picked it for his mother.
(2) Claire wished she had a pavif to add to the bouquet but all of hers were wilted.
(3) Tim tried to ruzel William during their fight but his fist met the wall instead.
(4) The teachers sometimes ruzel their students with rulers when they misbehave.
Each participant thus saw ten novel nouns and ten noun verbs, but each word was used in two different frames for a total of 40 experimental items per participant. As indicated in the examples above, the novel word had the same category meaning in both frames, and participants saw only one novel word for each noun or verb category. Across all participants, all novel words were read in both noun frames and both verb frames. The addition of the repetition manipulation resulted in a 2 (part of speech: noun or verb) x 2 (repetition condition: same-frame or different-frames) final design for the post-test data. For the eye movement data, the fact the novel words and sometimes the frames were read twice meant that the final design was 2 (part of speech: noun or verb) x 2 (repetition condition: same-frame or different-frames) x 2 (exposure: first or second).

**Apparatus.** The apparatus was the same as in Experiment 1.

**Procedure.** The procedure for the eye-tracking portion was the same as in Experiment 1. Participants read 40 experimental sentences and 40 fillers. 82.5% of the fillers were followed with yes-or-no comprehension questions to ensure that participants were staying on task. The average accuracy on these questions was 95%. Note that the first and second occurrence of each novel word were not in separate blocks, and there could therefore be anywhere from zero to 78 intervening sentences.

After the eye-tracking portion, the post-tests were administered. The familiarity rating task, part of speech test, category test, and their instructions were the same as in Experiment 1. After completing the post-test, participants completed the vocabulary subsection of the Nelson-Denny Reading Test, a standardized test of vocabulary knowledge (Brown, Fishco, & Hanna, 1993). The vocabulary test was followed by a reading span task to assess verbal working memory (Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005). In this task, participants
read sentences that were each followed by a letter. They read each sentence out loud and judged whether the sentence was realistic. After every block of two to five sentences, they were asked to recall the letters that had followed the sentences in the preceding block.

Results

Sentence regions and eye movement measures were the same as in Experiment 1. Duration measures were analyzed using linear mixed effect models, and binomial measures using logit mixed models. The models included part of speech, repetition condition, and exposure (whether the item was first or second occurrence of the target) as fixed effects, and random intercepts for participants and items. The three-way interaction of part of speech, repetition condition, and exposure was tested as a fixed effect for all measures but never significantly improved model fit and was thus removed for ease of interpretation. Interactions of exposure with part of speech or repetition condition were included when they improved model fit.

Target region. All seven eye movement measures were analyzed for the target region. Exposure had a main effect on all measures except regressions in: first fixation duration \((b = -44.08, SE = 11.63, t = -3.79)\); gaze duration \((b = -75.49, SE = 14.03, t = -5.38)\); go-past \((b = -101.53, SE = 15.07, t = -6.74)\); total time \((b = -171.98, SE = 22.31, t = -7.71)\); regressions out \((b = .95, SE = .24, z = -3.90, p < .0001)\); and skipping \((b = .87, SE = .20, z = 4.32, p < .0001)\). Targets were read for less time, less likely to be regressed out of, and more likely to be skipped when they were being seen for the second time.

The effect of exposure interacted with part of speech for first fixation duration \((b = 34.59, SE = 11.71, t = 2.95)\). When targets were read for the first time, nouns had longer first fixations
than verbs, but when read for the second time, verbs had longer durations than nouns. There were also interaction effects for regressions out ($b = .54, SE = .26, z = 2.10, p = .035$), and skipping ($b = .75, SE = .28, z = -2.65, p = .008$). The decrease in duration and likelihood of regressing, and the increase in skipping, that occurred when the target was repeated was larger for nouns than verbs for these measures.

Exposure interacted with repetition condition for first fixation duration ($b = -27.91, SE = 11.67, t = -2.37$) and regressions in ($b = 0.83, SE = .24, z = -3.48, p = 0.0005$). For these measures, the decrease in first fixation duration and likelihood of regressing that occurred when the target was repeated was larger when both targets were in the same sentence frame than when they were in different sentence frames.

Part of speech had an effect on first fixation duration ($b = -17.24, SE = 8.50, t = -2.03$), gaze duration ($b = 26.71, SE = 10.09, t = 2.65$), go-past time ($b = 36.35, SE = 10.78, t = 3.37$), total time ($b = 73.19, SE = 16.03, t = 4.57$), and regressions in ($b = .60, SE = .17, z = 3.54, p = .0003$). Verbs had longer reading times and were regressed into more often than nouns.

Repetition condition had an effect on go-past time ($b = -69.95, SE = 34.05, t = -2.05$) and total time ($b = -114.17, SE = 50.92, t = -2.24$). Targets had longer reading times when they were repeated in a different sentence frame than in the same sentence frame.

Part of speech and repetition condition interacted significantly for only one measure, regressions in ($b = .48, SE = .24, z = -1.99, p = 0.046$). Verbs repeated in a different frame were more likely to be regressed into than verbs repeated in the same frame and nouns in either frame.

**Pre-target region.** In the pre-target region, there was an effect of part of speech for total time ($b = 47.65, SE = 15.26, t = 3.12$) and regressions in ($b = .51, SE = .17, z = 2.98, p = .003$). Words preceding verb targets had longer reading times and more regressions in than words
preceding nouns. There was also an effect of exposure for total time \((b = -50.80, SE = 12.26, t = -4.14)\) and regressions in \((b = -0.62, SE = 0.12, z = -5.02, p < 0.0001)\). Reading times were shorter and fewer regressions in were made when the target was seen the second time.

**Post-target region.** Part of speech had an effect on total time \((b = 121.77, SE = 18.24, t = 6.68)\) and go-past time \((b = 105.25, SE = 13.77, t = 7.65)\). Post-target regions following a verb target had longer reading times than regions following noun targets. Part of speech interacted with exposure for total time \((b = -75.95, SE = 23.78, t = -3.19)\), go-past time \((b = -71.00, SE = 17.96, t = -3.95)\), and regressions out \((b = -0.76, SE = 0.34, z = -2.27, p = .023)\). The decrease in reading time and likelihood of regressing out that occurred when a target was seen a second time, versus the first time, was larger for post-target regions following verbs than following nouns. Exposure interacted with repetition condition for regressions out \((b = -1.24, SE = 0.43, z = -2.898, p = .004)\), such that the decrease in likelihood of regressing out of the post-target region that occurred when a target was seen the second time was larger when the target was in the same sentence frame.

There was an interaction between part of speech and repetition condition for regressions out as well as a three-way interaction between part of speech, repetition condition, and exposure; however, this effect was clearly drive by a difference between the same and different repetition conditions in the first exposure to the target. The repetition condition should have no effect on the first sentence, because at the time of reading the first sentence, the reader does not know whether the second sentence will be the same or different. Because the only possible explanation for this finding would be pre-cognition, these effects must be spurious.

**Sentence-ending region.** In the sentence-ending region, there was an effect of repetition condition on total time \((b = 173.68, SE = 51.76, t = 3.36)\) and regressions out \((b = 0.78, SE = 0.35, t = 2.27, p = .027)\).
$z = 2.24, p = .025$). Repetition condition and exposure also interacted for total time ($b = 298.69$, $SE = 58.44, t = -5.11$) and regressions out ($b = -.76, SE = .22, z = -3.52, p = .0004$). Both of these effects reflect that sentence endings had longer reading times and more regressions out when the target was repeated in a different frame instead of the same frame.

**Post-tests.** Accuracy on the part of speech post-test was better than chance for nouns in both repetition conditions (same: $t = 2.66, df = 209, p = .008$; different: $t = 4.65, df = 219, p < .0001$) and for verbs in the same-frame condition ($t = 2.09, df = 209, p = .038$), but not for verbs in the different-frames condition ($t = 0.878, df = 218, p = 0.3809$). Accuracy for nouns was better than verbs ($b = -.61, SE = .21, z = -2.94, p = .003$). There was a marginal interaction effect between part of speech and repetition condition ($b = .49, SE = .29, z = 1.71, p = .087$) reflecting that reading two different sentence frames instead of one frame twice improved accuracy for nouns but not for verbs. There was also an effect of vocabulary knowledge on accuracy ($b = .01, SE = .006, z = 2.02, p = .044$). Participants who scored higher on the Nelson-Denny vocabulary test were more accurate on the part of speech post-test.

Accuracy on the category post-test was better than chance for nouns in both repetition conditions (same: $t = 2.37, df = 209, p = 0.019$; different: $t = 4.65, df = 219, p < .0001$). The accuracy for verbs was only marginally better than chance in both conditions (same: $t = 1.80, df = 209, p = .07$; different: $t = 1.76, df = 219, p = 0.080$): however, noun accuracy was not significantly better than verb accuracy ($b = -.25, SE = .21, z = -1.23, p = .22$).

The familiarity ratings did not predict accuracy on either post-test, and part of speech did not have an effect on ratings. Ratings were affected by repetition condition ($b = -.82, SE = .36, z = -2.26, p = .024$), such that targets that had been seen in two different sentences frames were more likely to be rated familiar than those seen in the same sentence twice. There was an effect
of target length ($b = -0.18, \ SE = 0.06, z = -2.90, p = .003$) and an effect of inter-item distance ($b = -0.01, \ SE = 0.005, z = -2.219, p = .026$). Short targets were more likely to be rated familiar than long targets, and targets read in two sentences that occurred close together were rated as more familiar than targets read in two sentences separated by many items. Reading span also predicted rating ($b = .08, \ SE = .03, z = 3.04, p = .002$). Participants with larger reading spans rated more targets as familiar than those with smaller spans.

As in Experiment 1, mean-centered gazed durations and total times for the various sentence regions were tested as possible predictors for the two test accuracies and the familiarity ratings. Times on the first and second sentence were entered separately. One eye movement measure had an effect on part of speech accuracy: Longer total time on the target in the second sentence predicted higher accuracy ($b = .0007, \ SE = .0003, z = 2.67, p = .008$). The average total time for inaccurately answered targets was 340ms, whereas for accurately answered targets it was 391ms. The rest of the measures had no effect on the dependent variables and thus were not included in the final model fits.

**Discussion**

Experiment 2 found that reading times for novel nouns and verbs and their surrounding contexts decreased when they were read for a second time. This effect was particularly apparent in the likelihood of regressing. As predicted, this decrease was larger when reading the same sentences twice, but the fact that it occurred even when the words were in new sentences indicates that at the lowest level, readers recognized having seen the novel words before.

In the first experiment, there was a very consistent interaction effect between part of speech and novelty, part of which was the fact that novel noun targets were read for more time
than novel verb targets. Given that the sentences in the two experiments were mostly (though not exactly) identical, the reading times on the targets were expected to follow the same pattern the first time they were seen. This finding was not replicated. In fact, the opposite was true: Novel verb targets were read longer than novel nouns. The most likely explanation for this discrepancy is that readers in the second experiment adopted a strategy not used by readers in the first experiment. In Experiment 1, only a quarter of the sentences that participants read contained novel words: In Experiment 2, half of the total sentences contained novel words. In addition to this increase in proportion, participants in Experiment 2 saw novel words and even entire sentences repeated. Whereas Experiment 1 aimed to avoid drawing attention to the presence of novel words, that the novel words were central to the experiment was undoubtedly obvious in Experiment 2. It is unclear why this would have provoked longer reading times for verbs. Perhaps people in Experiment 1 exerted only the minimum effort needed to understand the contexts, while those in Experiment 2 were consciously trying to learn the words and had to exert much more effort on verbs because their meanings are harder to derive. This speculation is partially supported by the finding that novel nouns in this experiment still had longer first fixation durations than verbs in the first sentence: This earliest measure may have been less affected by the adoption of a conscious strategy.

Although reading times on the targets themselves differed, the two experiments showed similar patterns in readers’ use of context. In general, the words immediately before and after verbs were read more than nouns, as would be expected if unknown verbs are understood primarily as a relation between arguments. Less straightforward is the finding that the change in post-target reading time between the first and second sentence is larger for verbs than nouns. This effect appears to be driven by how inflated the reading times are for verb post-targets in the
first sentence. Because readers do not increase their reading of the context for nouns as much when they are first encountered, repetition has less of an effect. Similarly, the change in target reading between the first and second sentence is larger for nouns, because that is where extra effort is focused in the first case.

Repeating the novel targets had the intended effect of improving accuracy and exposing part of speech differences. Readers remembered part of speech better for nouns than verbs. Numerically, readers also remembered semantic category better for nouns – showing some learning for nouns but essentially chance performance for verbs – but they were not significantly different from each other. The effect of repetition condition on accuracy was unexpected and intriguing. As predicted, seeing two different contexts was more helpful than seeing the same context twice for nouns. For the verbs, seeing two different contexts actually hurt performance. Verbs appear to be more difficult to learn from context, so why is having another context, and thus more information, a hindrance? Perhaps learners are deriving verb meaning incorrectly in the first sentence, resulting in confusion when they get to the second sentence and the meaning does not fit. In this situation, I would expect reading times for verbs and their contexts in the second, different frame to be longer than they are, unless the confusion is so great that readers do not try to reconcile it.

A more likely explanation is that because novel verbs are interpreted in terms of their relation to the context, the meanings formed for them are more contextually bound than those of nouns. Kersten and Earles (2004) found a very similar phenomenon for memory of known nouns and verbs. In their experiments, participants were presented with simple intransitive sentences and instructed to remember either the nouns or the verbs. Afterwards, they were presented with some sentences that were identical to those seen earlier, some in which one of the words had
been seen earlier but the other was new, and some that were entirely new and asked to identify whether they had seen the nouns or verbs (whichever they had been told to remember) in the first part of the experiment. People were better at remembering the old verbs when they were paired with the old nouns than with new nouns, an effect that was much smaller for people trying to remember nouns. Kersten and Earles argue that memory for verbs is more contextually bound because verb semantics vary more across contexts than nouns (Gentner, 1981). This theory would in fact predict that verb learning would be better when the verbs were seen in the same context.
Chapter IV

General Discussion

The aim of the present study was to determine whether part of speech affects the ability to incidentally learn new words from written context. Although nouns and verbs typically differ in aspects like concreteness, which are known to affect learning, I attempted to reduce the influence of these external factors in order to isolate the effect of word class. In the learning tasks used here, there was clear advantage for nouns. Very little verb learning was found in either experiment. In addition to more traditional pen-and-paper knowledge assessments, the present study also took advantage of eye-tracking technology to better examine how novel words affect reading on-line. The reading measures collected for words and their contexts offer valuable insight into how the process of deriving word meaning differs for nouns and verbs. When a novel word is a noun, readers focus on the word itself, but when it is a verb, readers spend more time reading the context.

Context informativeness was not the source of these differences. I did not specifically control or quantify the number of available context clues for each word, but the results of the materials norming sub-experiment showed that readers were perfectly able to figure out the intended meanings of both nouns and verbs when they were explicitly asked. Something that grows and can be picked was interpreted as probably some kind of flower just as something done with a fist during a fight was interpreted as probably some manner of hitting. When readers were asked only to read and comprehend the contexts, the processes for interpreting sentences containing unknown nouns and unknown verbs diverged. Readers appeared to quickly extract the necessary contextual information to identify the nouns and were at least somewhat able to remember that information later. The amount of time readers spent reading the context for verbs,
on the other hand, suggests that they attempted to establish the relational framework but may not have fully abstracted the meaning of the verb from that relationship information. Reading words in different contexts may not have been facilitative because readers were not decontextualizing the verb meanings, making it difficult to combine meanings from two sentences.

These two distinct approaches to processing unknown words in text reflect the semantic and syntactic differences between nouns and verbs, discussed throughout this report. Many nouns refer to objects that are perceived and even categorized similarly across languages (Gentner, 1981; Rosch et al., 1976). Noun semantics are internally dense and stable across contexts. Verbs in general do not have this same cohesion or stability: Their meanings are more variable and often context dependent. Context strongly influences how a verb is interpreted (Gentner & France, 1988), and even how well it is remembered (Kersten & Earles, 2004). The reading and learning results of the present study are consistent with a wealth of evidence, spanning many areas of research, showing that verbs are just harder – harder to define, harder to process, harder to remember, and harder to learn.

**Limitations and Future Directions**

This study was designed to compare incidental learning of nouns and verbs. The primary goal in constructing the sentence contexts was to make the context following the target word highly informative. Some guidelines were used to make them conducive for eye-tracking, such as not placing targets too near the beginning of the sentence and keeping the sentences all on one line, but the sentences were to avoid appearing formulaic. This method meant that the items were well suited for comparing reading of real and novel words, but not for comparing reading of nouns and verbs. The verb targets always had direct objects, but the syntactic structure of the
sentence otherwise varied. The nouns were usually, but not always, direct objects. The verbs were sometimes infinitives and not the main verb. There was little consistency in the syntactic or semantic role of the pre- and post-target regions. The sentence-ending regions sometimes contained separate clauses. My ability to draw specific conclusions about the part of speech differences in the reading measures is thus rather limited. However, the fact that consistent patterns for nouns and verbs emerged despite all of the variation is strong evidence that there are interesting differences in how people go about deriving meanings for these words. Studying context use more systematically using structurally uniform sentences would elucidate the nature of these differences, and would be an ideal extension of this work.

Previous research has shown that it is possible to gain measurable knowledge about a new word from a single exposure in context (Borovsky et al., 2010; Chaffin et al., 2001; Diakidoy, 1998; Nagy et al., 1987; Nagy et al., 1985; Williams & Morris, 2004). Williams and Morris (2004), whose method was closely mirrored in the present study, found that readers learned new words well enough to select their correct category. Unfortunately, I found very little evidence of learning. The second experiment was designed specifically to result in more learning, and while learning did improve, the gains were modest. In the condition with the best performance, nouns used in two different contexts, the average number of correct answers on the easier part of speech test was 6.5 – guessing would have resulted in an average of five correct answers.

In most regards, learning in this study should have been easy compared to learning from natural prose. The words were all highly imageable. They were, by Nagy et al.’s (1987) definition, as conceptually simple as possible: synonyms for known concepts. The contexts were highly informative: readers who were directly asked what should fit in the space occupied by the
novel word were able to do so. Participants were tested on the words immediately after reading them so decay was minimized. On the other hand, the brevity of the contexts and the concentration of new words probably contributed to the difficulty of learning. For whatever reason, one and two exposures were not enough for this set of readers to gain much word knowledge from this set of materials. Future research could examine the part of speech effect using testing measures that are more sensitive to subtle degrees of knowledge, such as those tested by Durso and Shore (1991), to better understand the differences occurring at the earliest point of learning. Noun and verb learning could also be compared using a technique designed to measure small incremental changes in word meaning representations over many exposures, such as that presented by Frishkoff et al. (2008). Such a comparison would produce more fine-grained data on both the speed and the nature of incremental word learning. In light of the present study’s findings regarding the effects of reading a word in multiple contexts, it is possible that noun and verb representations follow different trajectories as they develop. This method could be used, for example, to test my conjecture about the failure to abstract and combine verb meaning information from multiple sources.

In the present study, readings measures rarely predicted performance on the post-tests. In the first experiment, accuracy may have been too low to detect any connection with eye movements. In the second experiment, spending more overall time on the target in second context was related with better accuracy, but only on the part of speech test. It is unclear why extra effort at the latest possible opportunity would improve memory for part of speech but not category. This is unlike the relationship between reading and accuracy reported in Williams and Morris (2004). They found that novel words whose meanings were correctly selected on the post-test were read more quickly initially but re-read for more time than words whose meanings
were not correctly selected. They suggested that words might be learned better when they are quickly identified as novel and their contexts referenced before trying to process the words themselves. While I did not find this pattern myself, the results of the present study do not permit me to comment on their interpretation: Experiment 1 lacked the power to detect such a pattern, and Experiment 2 was too unlike their method. It is somewhat surprising that reading time was not more predictive of accuracy. Future research should continue to probe this relationship in order to learn what sort of reading behavior contributes to learning success.

This report does not exhaust the possible investigations that could be made with the data collected from the present experiments. For instance, the length of the novel word was a strong predictor of whether the word would be judged as familiar in both experiments. Although target word length did not affect accuracy, it is possible that reading patterns on short and long words differed. Similarly, verbal ability and working memory were not tested as predictors of reading time. These analyses were deemed to be outside the scope of the present study, but they could be easily conducted as part of continuing work on this topic. Their results may suggest other fruitful directions for future research.

Conclusion

This study examined the effect of part of speech on incidental word learning from one and two exposures. Overall, very little learning was demonstrated across the study, but it was still possible to determine that nouns were generally learned better than verbs. Reading times on the sentences contexts showed distinct patterns for novel nouns and novel verbs. Novel verbs tended to elicit more extensive use of contextual information than novel nouns. Both the learning and the eye movement results were consistent with claims from various areas of cognitive
science that verbs are less semantically cohesive, more contextually dependent, and generally harder to learn.

Incidental word learning is not easy to observe or to measure. Research on the topic has amassed information on how factors at the learner-, context, and word-level effect learning success, but the nature of the learning process itself remains elusive. While people use explicit clues from the context to figure out the meanings of unknown words when prompted to determine an unknown word’s meaning, there is less evidence that readers engage in such behavior when reading for other purposes. Yet, somehow, readers do manage to learn words during normal reading. Traditional pen-and-paper tests may be approaching the end of their usefulness for studying contextual word learning. Future research on incidental word learning will need to incorporate measures of on-line mental processing, such as eye movements and ERPs, if the process is ever to be understood.
References


doi:10.3758/BRM.40.4.907


doi:10.1080/10888438.2011.539076


# Tables and Figures

## Table 1

**Target Words**

<table>
<thead>
<tr>
<th>Group</th>
<th>Noun</th>
<th>Noun category</th>
<th>Verb</th>
<th>Verb category</th>
<th>Novel Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>pigeon</td>
<td>bird</td>
<td>devour</td>
<td>eat</td>
<td>muloid</td>
</tr>
<tr>
<td>A</td>
<td>lung</td>
<td>body part</td>
<td>chop</td>
<td>cut</td>
<td>vorl</td>
</tr>
<tr>
<td>A</td>
<td>belt</td>
<td>clothing</td>
<td>toss</td>
<td>throw</td>
<td>kuwn</td>
</tr>
<tr>
<td>A</td>
<td>daisy</td>
<td>flower</td>
<td>smack</td>
<td>hit</td>
<td>pavif</td>
</tr>
<tr>
<td>A</td>
<td>mango</td>
<td>fruit</td>
<td>grill</td>
<td>cook</td>
<td>ikift</td>
</tr>
<tr>
<td>A</td>
<td>diamond</td>
<td>jewel</td>
<td>inspect</td>
<td>examine</td>
<td>tirlask</td>
</tr>
<tr>
<td>A</td>
<td>piano</td>
<td>musical instrument</td>
<td>crack</td>
<td>break</td>
<td>treve</td>
</tr>
<tr>
<td>A</td>
<td>golf</td>
<td>sport</td>
<td>wash</td>
<td>clean</td>
<td>udat</td>
</tr>
<tr>
<td>A</td>
<td>carrot</td>
<td>vegetable</td>
<td>clutch</td>
<td>hold</td>
<td>sagoll</td>
</tr>
<tr>
<td>A</td>
<td>bicycle</td>
<td>vehicle</td>
<td>conceal</td>
<td>hide</td>
<td>pruiche</td>
</tr>
<tr>
<td>B</td>
<td>hawk</td>
<td>bird</td>
<td>dice</td>
<td>cut</td>
<td>luit</td>
</tr>
<tr>
<td>B</td>
<td>ankle</td>
<td>body part</td>
<td>punch</td>
<td>hit</td>
<td>ruzel</td>
</tr>
<tr>
<td>B</td>
<td>sweater</td>
<td>clothing</td>
<td>shatter</td>
<td>break</td>
<td>dasponk</td>
</tr>
<tr>
<td>B</td>
<td>tulip</td>
<td>flower</td>
<td>steam</td>
<td>cook</td>
<td>betik</td>
</tr>
<tr>
<td>B</td>
<td>pear</td>
<td>fruit</td>
<td>mask</td>
<td>hide</td>
<td>ehup</td>
</tr>
<tr>
<td>B</td>
<td>ruby</td>
<td>jewel</td>
<td>grip</td>
<td>hold</td>
<td>pehl</td>
</tr>
<tr>
<td>B</td>
<td>accordion</td>
<td>musical instrument</td>
<td>disinfect</td>
<td>clean</td>
<td>narsomple</td>
</tr>
<tr>
<td>B</td>
<td>karate</td>
<td>sport</td>
<td>nibble</td>
<td>eat</td>
<td>zerden</td>
</tr>
<tr>
<td>B</td>
<td>potato</td>
<td>vegetable</td>
<td>review</td>
<td>examine</td>
<td>halcar</td>
</tr>
<tr>
<td>B</td>
<td>yacht</td>
<td>vehicle</td>
<td>fling</td>
<td>throw</td>
<td>joler</td>
</tr>
</tbody>
</table>
Table 2

*Mean Reading Times on the Target Region by Novelty and Part of Speech*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Real Noun</th>
<th>Real Verb</th>
<th>Novel Noun</th>
<th>Novel Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>First fixation duration</td>
<td>187 (114)</td>
<td>213 (117)</td>
<td>264 (124)</td>
<td>257 (136)</td>
</tr>
<tr>
<td>Gaze duration</td>
<td>218 (170)</td>
<td>252 (174)</td>
<td>424 (303)</td>
<td>392 (269)</td>
</tr>
<tr>
<td>Go-past time</td>
<td>227 (175)</td>
<td>264 (183)</td>
<td>492 (331)</td>
<td>477 (318)</td>
</tr>
<tr>
<td>Total time</td>
<td>270 (236)</td>
<td>341 (284)</td>
<td>676 (470)</td>
<td>677 (479)</td>
</tr>
<tr>
<td>Regressions in</td>
<td>.15 (.36)</td>
<td>.22 (.41)</td>
<td>.38 (.49)</td>
<td>.36 (.48)</td>
</tr>
<tr>
<td>Regressions out</td>
<td>.18 (.39)</td>
<td>.18 (.38)</td>
<td>.26 (.44)</td>
<td>.30 (.46)</td>
</tr>
<tr>
<td>Skipping</td>
<td>.27 (.44)</td>
<td>.21 (.41)</td>
<td>.10 (.30)</td>
<td>.12 (.33)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations are in parentheses.
Table 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>First fixation</th>
<th>Gaze duration</th>
<th>Go-past time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Novelty</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>-77.45</td>
<td>-204.49</td>
<td>-265.90</td>
<td>-406.76</td>
</tr>
<tr>
<td>$SE$</td>
<td>9.36</td>
<td>16.82</td>
<td>17.85</td>
<td>25.58</td>
</tr>
<tr>
<td>$t$</td>
<td>-8.27</td>
<td>-12.16</td>
<td>25.58</td>
<td>-15.90</td>
</tr>
<tr>
<td><strong>Part of speech</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>-6.86</td>
<td>-33.23</td>
<td>-19.53</td>
<td>-8.83</td>
</tr>
<tr>
<td>$SE$</td>
<td>10.88</td>
<td>22.29</td>
<td>24.31</td>
<td>32.79</td>
</tr>
<tr>
<td>$t$</td>
<td>-.63</td>
<td>-1.49</td>
<td>-.80</td>
<td>-.27</td>
</tr>
<tr>
<td><strong>Novelty x Part of speech</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>32.57</td>
<td>67.30</td>
<td>56.18</td>
<td>78.55</td>
</tr>
<tr>
<td>$SE$</td>
<td>13.24</td>
<td>23.78</td>
<td>25.21</td>
<td>36.13</td>
</tr>
<tr>
<td>$t$</td>
<td>2.46</td>
<td>2.83</td>
<td>2.23</td>
<td>2.17</td>
</tr>
</tbody>
</table>

Estimates, SEs, and t-Values of Fixed Effects for Target Region Reading Times

Note. Effects with absolute value of $t > 2$ are significant at $p = .05$. 
### Table 4

*Estimates, SEs, z-Values, and p-Values of Fixed Effects for Target Region Binomial Reading Measures*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Regressions in</th>
<th>Regressions out</th>
<th>Skipping</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Novelty</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>-1.35</td>
<td>-.51</td>
<td>1.30</td>
</tr>
<tr>
<td>$SE$</td>
<td>.20</td>
<td>.20</td>
<td>.23</td>
</tr>
<tr>
<td>$z$</td>
<td>-6.86</td>
<td>-2.55</td>
<td>5.65</td>
</tr>
<tr>
<td>$p$</td>
<td>&lt;.001</td>
<td>.01</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Part of speech</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>-.10</td>
<td>.18</td>
<td>.26</td>
</tr>
<tr>
<td>$SE$</td>
<td>.20</td>
<td>.22</td>
<td>.27</td>
</tr>
<tr>
<td>$z$</td>
<td>-.53</td>
<td>-.82</td>
<td>.98</td>
</tr>
<tr>
<td>$p$</td>
<td>.59</td>
<td>.41</td>
<td>.33</td>
</tr>
<tr>
<td><strong>Novelty x Part of speech</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>.54</td>
<td>-.20</td>
<td>-.63</td>
</tr>
<tr>
<td>$SE$</td>
<td>.27</td>
<td>.28</td>
<td>.32</td>
</tr>
<tr>
<td>$z$</td>
<td>2.02</td>
<td>-.73</td>
<td>-1.96</td>
</tr>
<tr>
<td>$p$</td>
<td>.04</td>
<td>.47</td>
<td>.05</td>
</tr>
</tbody>
</table>
Table 5

*Mean Reading Times on the Target Region by Exposure and Part of Speech*

<table>
<thead>
<tr>
<th>Variable</th>
<th>First sentence</th>
<th></th>
<th>Second sentence</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Noun</td>
<td>Verb</td>
<td>Noun</td>
<td>Verb</td>
</tr>
<tr>
<td>First fixation duration</td>
<td>266 (144)</td>
<td>249 (115)</td>
<td>216 (132)</td>
<td>234 (129)</td>
</tr>
<tr>
<td>Gaze duration</td>
<td>364 (238)</td>
<td>373 (254)</td>
<td>263 (193)</td>
<td>303 (218)</td>
</tr>
<tr>
<td>Go-past time</td>
<td>406 (263)</td>
<td>426 (298)</td>
<td>277 (205)</td>
<td>324 (228)</td>
</tr>
<tr>
<td>Total time</td>
<td>546 (385)</td>
<td>620 (479)</td>
<td>336 (282)</td>
<td>404 (318)</td>
</tr>
<tr>
<td>Regressions in</td>
<td>.30 (.46)</td>
<td>.38 (.49)</td>
<td>.17 (.38)</td>
<td>.22 (.42)</td>
</tr>
<tr>
<td>Regressions out</td>
<td>.25 (.43)</td>
<td>.26 (.44)</td>
<td>.13 (.34)</td>
<td>.20 (.40)</td>
</tr>
<tr>
<td>Skipping</td>
<td>.11 (.32)</td>
<td>.15 (.36)</td>
<td>.22 (.42)</td>
<td>.16 (.37)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations are in parentheses.
Table 6

*Mean Accuracy and Familiarity Rating by Repetition Condition and Part of Speech*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Same frame</th>
<th></th>
<th>Different frame</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Noun</td>
<td>Verb</td>
<td>Noun</td>
<td>Verb</td>
</tr>
<tr>
<td>Part of speech accuracy</td>
<td>.59 (.49)</td>
<td>.57 (.50)</td>
<td>.65 (.48)</td>
<td>.53 (.50)</td>
</tr>
<tr>
<td>Category accuracy</td>
<td>.58 (.50)</td>
<td>.56 (.50)</td>
<td>.63 (.48)</td>
<td>.56 (.50)</td>
</tr>
<tr>
<td>Familiarity ratings</td>
<td>.67 (.47)</td>
<td>.68 (.47)</td>
<td>.82 (.38)</td>
<td>.80 (.40)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviation are in parentheses.
<table>
<thead>
<tr>
<th>20 Real Nouns . . .</th>
<th>A: 10 real nouns ($A_{\text{real nouns}}$) and their frames ($A_{\text{noun frames}}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B: 10 real nouns ($B_{\text{real nouns}}$) and their frames ($B_{\text{noun frames}}$)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>20 Real Verbs . . .</th>
<th>A: 10 real verbs ($B_{\text{real verbs}}$) and their frames ($B_{\text{verb frames}}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B: 10 real verbs ($B_{\text{real verbs}}$) and their frames ($B_{\text{verb frames}}$)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>20 Novel Words . . .</th>
<th>A: 10 novel words ($A_{\text{novel words}}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B: 10 novel words ($B_{\text{novel words}}$)</td>
</tr>
</tbody>
</table>

---

**Experiment 1**

**List 1**

<table>
<thead>
<tr>
<th>Real Nouns</th>
<th>Noun Frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{\text{real nouns}}$</td>
<td>$A_{\text{noun frames}}$</td>
</tr>
<tr>
<td>$B_{\text{real nouns}}$</td>
<td>$B_{\text{noun frames}}$</td>
</tr>
</tbody>
</table>

**List 2**

<table>
<thead>
<tr>
<th>Real Verbs</th>
<th>Verb Frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_{\text{real verbs}}$</td>
<td>$B_{\text{verb frames}}$</td>
</tr>
</tbody>
</table>

**Experiment 1**

*Note.* The top half shows the division of targets into groups A and B. The bottom half shows how novelty and part of speech were crossed and counterbalanced across two lists. Groups A and B correspond with groups A and B in Table 1.

*Figure 1.* Design of Experiment 1
### Experiment 2

**Same Repetition Condition**

<table>
<thead>
<tr>
<th>List 1</th>
<th>List 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{\text{novel words}} + A_{\text{noun frames}}$</td>
<td>$B_{\text{novel words}} + B_{\text{noun frames}}$</td>
</tr>
<tr>
<td>$B_{\text{novel words}} + B_{\text{verb frames}}$</td>
<td>$A_{\text{novel words}} + A_{\text{verb frames}}$</td>
</tr>
<tr>
<td>$\times 2$</td>
<td>$\times 2$</td>
</tr>
</tbody>
</table>

### Experiment 2

**Different Repetition Condition**

<table>
<thead>
<tr>
<th>List 1</th>
<th>List 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{\text{novel words}} + A_{\text{noun frames}}$</td>
<td>$B_{\text{novel words}} + B_{\text{noun frames}}$</td>
</tr>
<tr>
<td>$+ B_{\text{noun frames}}$</td>
<td>$+ A_{\text{noun frames}}$</td>
</tr>
<tr>
<td>$B_{\text{novel words}} + B_{\text{verb frames}}$</td>
<td>$A_{\text{novel words}} + A_{\text{verb frames}}$</td>
</tr>
<tr>
<td>$+ A_{\text{verb frames}}$</td>
<td>$+ B_{\text{verb frames}}$</td>
</tr>
<tr>
<td>$\times 2$</td>
<td>$\times 2$</td>
</tr>
</tbody>
</table>

*Note.* The top half shows how novelty was crossed and counterbalanced across two lists in the same-frame repletion condition. $X2$ indicates that these items occurred twice. The bottom half shows how novelty was crossed and counterbalanced across two lists in the different-frames repetition condition. In this case, two frame groups were used with each novel word group.

*Figure 2.* Design of Experiment 2
Appendix

The sentence contexts used in the reading tasks are presented. Participants viewed the frames either with a real target word, as they are shown here, or with a novel target. The real word targets are italicized for emphasis, but were not italicized during the experiments.

Noun frames

Nora always enjoyed *golf* but she had to quit playing after she hurt her back.

Noah took the *accordion* from its case and began to play us a traditional polka.

Eric hoped the *mango* he had left was not too ripe for him to use in a smoothie.

Claire wished she had a *tulip* to add to the bouquet but all of hers were wilted.

Tom learned that his *lung* has a tumor in it that will have to be removed soon.

Constance witnessed a *hawk* swooping down from the sky to catch a field mouse.

Jason saw a *daisy* growing near the mailbox and he picked it for his mother.

Scott hoped that his *ankle* wouldn't get too swollen after he slipped and fell.

Rebecca took the *sweater* from her closet and tried it on in front of the mirror.

My family owned a *piano* when I was young and my mother used to play it for me.

There was a large *ruby* set in the crown that the museum added to its exhibit.

Nathan grabbed a *carrot* from the refrigerator and sliced it up for his salad.

Dan is learning *karate* and hopes to compete in the big tournament next month.

I hate the *pigeon* that built a nest on my windowsill because he coos all day.

Larry has a *yacht* and will take us for a ride when we visit his beach house.

Carl needs a new *belt* because he ripped the buckle off his old one yesterday.

Max hopes the *potato* he planted in his garden will grow large enough to cook.
Jane couldn't believe the diamond in the ring her boyfriend gave her was so big.
Emily enjoyed the pear she had packed in her lunch because it was very juicy.
Brian said that his bicycle broke on his way to work and asked for a ride home.

**Verb frames**

Joseph decided to wash his car before going to pick up his date for the dance.
Megan's job is to disinfect the tools that the doctors need to perform surgery.
We want to grill the chicken for the picnic tomorrow but we are out of charcoal.
Alice likes to steam the broccoli until tender before adding it to the dish.
Louis went to chop some wood for the fire but he broke his axe on a thick log.
Her two assistants dice the vegetables very quickly with their sharp knives.
The teachers sometimes smack their students with rulers when they misbehave.
Tim tried to punch William during their fight but his fist met the wall instead.
The kids might shatter one of Sue's windows while playing baseball in the yard.
My neighbors always crack the ice covering their pond so we can’t skate on it.
My grandmother will grip the handrail tightly whenever she goes down the stairs.
The women tourists clutch their purses as they walk through the crowded plaza.
The mice must nibble the crumbs quickly or the homeowner will discover them.
Cindy's teenage sons can devour a week's worth of groceries in a single meal.
Gabe will be able to fling objects very far if he practices using his slingshot.
Shawn didn't want to toss the ball to me because he knew I wouldn't catch it.
The manager will review his resume to decide if he is qualified to work here.
The police officers inspect the crime scene closely before submitting a report.
Ian tried to *mask* the smell of sweat with cologne but it wasn't strong enough.

Makeup artists can *conceal* blemishes and scars but the cover is only temporary.