Published in final edited form as:

J Cogn Psychol (Hove). 2012 ; 24(2): 179–188. doi:10.1080/20445911.2011.604028.

Lexical embeddings produce interference when they are morphologically unrelated to the words in which they are contained: Evidence from eye movements

Kristin M. Weingartner,
Hofstra University

Barbara J. Juhasz, and
Wesleyan University

Keith Rayner
University of California, San Diego

Abstract

Many words in the English language contain semantically and morphologically unrelated smaller words (e.g., *room* in *groom*). Recent findings indicate that a high frequency embedded word produces interference during visual word identification (e.g., Bowers, Davis, & Hanley, 2005; Davis, Perea, & Acha, 2009; Davis & Taft, 2005). In an eye movement experiment we examined whether lexical embeddings produce interference even when explicit judgments about lexicality or category membership are not solicited. Participants silently read sentences that each contained a target word with a lexical embedding. Fixation times were longer on target words that contained a higher frequency embedding compared to those that contained a lower frequency embedding. This finding indicates that a high frequency embedding interferes with word identification during silent reading and adds to a growing body of evidence that a word's orthographic neighborhood includes embedded words.

Keywords

Reading; word recognition; eye movements; orthographic neighbor; lexical embedding

During reading, a variety of processes unfold which allow a reader to understand the text currently under inspection. One important question in the field is how individual words are recognized and understood. This question is more complex than it might appear to be on the surface because the term “word recognition” may be used to refer to a variety of processes including access to the orthographic word form, access to the phonology of the word, and access to the meaning of the word.

There currently exist several models of visual word recognition which differ from each other in terms of general architecture. With many of these models word representations are realized in terms of an orthographic lexicon module (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) while others rely on distributed orthographic representations (e.g., Harm & Seidenberg, 2004). Thus, these modeling efforts support the idea that the orthography of words is important in processing during reading. Empirical findings from investigations with

skilled readers provide evidence that orthographic similarity influences how words are processed. This is clearly shown in studies which examine the role of a word's orthographic neighborhood. A word's orthographic neighborhood size (N) is traditionally defined as the number of words that differ from the base word by a single letter (Coltheart, Davelaar, Jonasson, & Besner, 1977). Thus, the orthographic neighbors for the word *dog* would be *dot*, *log*, *dig*, and so on. Most studies have shown facilitation for words with a large neighborhood when compared to words with a small neighborhood and this finding holds in both lexical decision and naming (Andrews, 1997).

However, orthographic similarity does not always facilitate word recognition. When overall neighborhood size is controlled, the majority of studies have demonstrated that having a higher frequency orthographic neighbor produces inhibition in the lexical decision task (Andrews, 1997). Notably, there appear to be some language differences with this neighborhood frequency effect. Lexical decision experiments in Spanish and French have consistently shown an inhibitory neighborhood frequency effect while some studies in English have demonstrated null effects or facilitation from this variable (for a review see Andrews, 1997). The role of neighbor frequency has been less extensively studied in the word naming task, but 3 out of the 4 studies included in the Andrews (1997) review showed facilitation for having a higher frequency neighbor.

The role of a word's orthographic neighborhood has also been investigated in eye movement experiments (Paterson, Liversedge, & Davis, 2009; Perea & Pollatsek, 1998; Pollatsek, Perea, & Binder, 1999; Slattery, 2009; Williams, Perea, Pollatsek, & Rayner, 2006). Pollatsek et al. (1999) found that the facilitory effect of neighborhood size observed in lexical decision and word naming are not observed during more natural reading situations. They inserted words containing either a large or small neighborhood into sentences and recorded readers' eye movements. Gaze durations were significantly longer on words with many neighbors compared to words with few neighbors. This task difference is made even more striking in that Pollatsek et al. demonstrated that the same target words used in the eye-tracking study produced facilitation in a standard lexical decision task. Perea and Pollatsek (1998) also observed inhibition in an eye movement experiment when the presence of a higher frequency neighbor was manipulated. Here, the effects were not observed on the target word itself, but were observed on the post-target region and in the percentage of regressions back to the target word (however, cf. Sears, Campbell, & Lupker, 2006; but see also Slattery, 2009).

Each of the studies discussed above used a traditional definition of an orthographic neighborhood, where neighbors are defined as words differing by only a single letter. However, several other findings (Bowers, Davis, & Hanley, 2005; Davis, Perea, & Acha, 2009; Davis & Taft, 2005; McCormick, Davis, & Brysbaert, 2010; Perea, Acha, & Fraga, 2008) suggest that the concept of orthographic neighborhood needs to be extended to include more than just the traditional "substitution neighbors" described above. Of particular relevance to the current article, there is a growing body of evidence that "deletion neighbors", words that differ from their base word by the deletion of a letter (e.g., *last* in *blast*), are activated during word recognition. For example, with English (Davis & Taft, 2005) and Spanish (Davis et al., 2009), response times in the lexical decision task were longer for low frequency words that contained a higher frequency deletion neighbor relative to matched words that did not contain a deletion neighbor. Thus, consistent with the conclusions from several investigations of substitution neighbors, it appears as though having a higher frequency deletion neighbor produces inhibition in the lexical decision task.

The findings from some semantic categorization studies suggest that deletion neighbors may even be activated to the point of semantics in the mental lexicon (Bowers et al., 2005;

McCormick et al., 2010; Nation & Cocksey, 2009). In these studies a semantic category (e.g., food or clothing) was presented on a computer screen and participants had to respond as quickly and accurately as possible as to whether a target word was a member of the specified category. Each target word contained a smaller “embedded” word (*hat* in *hatch*). In the congruent condition, neither the target word nor the embedded word were members of the category. In the incongruent condition, the target word was not a member of the category, but the embedded word was a member (e.g., clothing: *hatch*). Thus, for both conditions the correct answer was that the target word was *not* a member of the semantic category. Response times for the categorization decisions were significantly longer in the incongruent condition, which suggests that the meaning of the embedded word was activated during the task and produced interference. This effect appears to hold only when the embedded word is of higher frequency than the target word (Bowers et al., 2005).

It is interesting to note that the position of an embedded word within its target has varied within and across investigations. The summed pattern of these results suggests that embedded words produce interference for target word processing across each of several positions: initial (*earn*; Bowers et al., 2005; Davis et al., 2009; Davis & Taft, 2005; McCormick et al., 2010), middle (*scowl*; Bowers et al., 2005), final (*howl*; Bowers et al., 2005; Davis & Taft, 2005; McCormick et al., 2010), and outer (*trance*; Davis et al., 2009; Davis & Taft, 2005).

Although there is a growing body of evidence that embedded words affect word recognition, these findings have been obtained largely from the lexical decision and semantic categorization tasks, each of which is a significant departure from natural reading. One exception from the Spanish language is Davis et al. (2009). They had participants read sentences with target words that either had a higher frequency embedded word in the outer position or did not (*alternar* has the higher frequency embedded word *alterar*). The mean gaze and total fixation durations were longer higher on targets that contained a neighbor. This finding suggests that the inhibitory effect of embedded words occurs during more natural reading situations, at least with Spanish.

The purpose of the present experiment was to examine the effects of embedded words during natural reading with the English language. While research has been devoted to this question in the past (e.g. Juhasz, 2007; Juhasz, Starr, Inhoff, & Placke, 2003; Niswander-Klement & Pollatsek, 2006; Niswander, Pollatsek, & Rayner, 2000), in these investigations the embedded word was the root or lexeme of a morphologically complex word and the presence of a high frequency root or lexeme led to *shorter* fixation durations on the morphologically complex word (for long prefixed words see Niswander-Klement & Pollatsek, 2006; for derived suffixed words see Niswander et al., 2000; for compound words see Juhasz, 2007 and Juhasz et al., 2003).

In the present experiment, we examined the effects of embedded words during reading with the English language. Although the results from Davis et al.’s (2009) investigation with Spanish sentences suggest that non-morphologically related embedded words interfere with word identification during silent reading, it is not clear that this conclusion will generalize to English for a number of reasons. First, as mentioned above, Andrews’ (1997) review of traditional orthographic neighborhood effects demonstrated that neighborhood frequency effects produce more reliable inhibition in Spanish in the lexical decision task compared to English. Second, even if an effect of embedded words is observed in English, it might be qualitatively different than that observed in lexical decision. The orthographic neighborhood size effects discussed previously illustrate this point. Having a large orthographic N facilitates lexical decision responses, but produces inhibitory effects during reading (Pollatsek et al., 1999).

In order to investigate the role of lexical embeddings, sentences were constructed in which a critical target word contained an unrelated embedded word that was either higher or lower in frequency than its target. For example, with the targets *scar* and *lamp*, *car* is a higher-frequency embedded word and *amp* is a lower-frequency embedded word. The full list of materials appears in the Appendix. Readers' eye movements were recorded as they read the sentences silently. Based on previous research (Bowers et al., 2005; Davis et al., 2009; Davis & Taft, 2005; McCormick et al., 2010) we expected higher frequency embedded words to have an inhibitory role on target word processing. This would lead to longer fixation times on targets containing higher frequency embedded words than targets containing lower frequency embedded words.

The present study extends the work of Davis et al. (2009) not only by examining eye movements on embedded words in a different language, but also by examining different embedding locations. The Davis et al. (2009) study utilized outer deletion neighbors while we examined embedded words in the initial and final positions. In addition, the baseline condition was different in the current study. Davis et al. (2009) used control words with no deletion neighbors. In the current study, the control words also contained an embedded word. Finally, the embedded words in the current study differed from their base word by the deletion of one to three letters (e.g. *man* in *manner*) and thus do not adhere to the definition of "deletion neighbor" in Davis et al.'s (2009) study.

Method

Participants

Thirty members of the University of Massachusetts community participated in exchange for either extra credit in a psychology course or for payment (\$8.00). All participants were native speakers of American English and had normal vision or wore contact lenses.

Apparatus

Eye movements were recorded using a Fourward Technologies (Buena Vista, VA) Dual Purkinje Eyetracker (Generation VI) interfaced with an IBM computer. The eye tracker has a resolution of less than 10 minutes of arc. Eye movements were measured from the right eye but viewing was binocular. Participants were seated 61 cm from the computer screen where the single line sentences were presented. At this distance 3.8 characters subtended 1 degree of visual angle. The sentences were displayed on a 15-inch NEC MutliSync 4FG monitor.

Procedure

After reading a brief description of the experimental procedure, participants were verbally instructed about the task. In order to minimize head movements during the experiment, a bite bar was prepared out of dental clay for each participant and a forehead rest was used. A calibration routine was performed, consisting of a 9-point horizontal calibration grid, and its accuracy was checked after each sentence and redone as deemed necessary. A Yes/No comprehension question followed 17% of the trials during the experiment. Participants responded to each question by pressing a corresponding Yes or No button on a response box. The average accuracy on the questions was above 97%.

Stimuli

The target words consisted of 36 nouns which were selected from the English Lexicon Project (Balota et al., 2007). Each target word contained an embedded noun that resulted from the deletion of one to three letters from its respective target word. Half of the targets

contained an embedded word that was higher in frequency (e.g., *army*) and the other half contained an embedded word that was lower in frequency (e.g., *firm*). The higher frequency embedded words had an average frequency of 310 per million and the lower frequency embedded words had an average frequency of 8 per million (Kuera & Francis, 1967), $t(34) = 3.78$, $SEM = 79.98$, $p = .001$.¹ As a manipulation check, lexical decision and naming latencies from the English Lexicon Project were analyzed for the 36 target words. These analyses confirmed that the higher frequency embedded words interfered with target identification, although the effect was reliable only in naming. The mean difference between conditions was 26 ms for lexical decision ($t(34) = 1.26$, $SEM = 20.76$, $p > .10$) and 36 ms for naming ($t(34) = 2.39$, $SEM = 14.91$, $p = .024$).

To provide the purest possible investigation of lexical embedding, the 18 targets containing a higher frequency embedded word were well matched with the 18 targets containing a lower frequency embedded word on several characteristics that have been shown to influence word identification. These characteristics are reported in Table 1. Information about word length, frequency, neighborhood size (i.e., number of substitution neighbors), and the number of higher-frequency substitution neighbors was obtained from the English Lexicon Project (Balota et al., 2007). Separate normative studies were conducted to ensure that the two groups of target words were well matched with respect to familiarity (12 raters), concreteness (12 raters), and age of acquisition (12 raters). All ratings used a 1–7 scale (for age-of-acquisition, the Gilhooly and Logie (1980) scale was used). Importantly, the two groups of target words did not differ significantly with respect to any of the characteristics illustrated in Table 1 (all $ps > .18$).

The 36 target words were divided into 18 pairs with the following constraints. First, for each pair, one target contained a higher frequency embedded word and the other contained a lower frequency embedded word. Second, the targets in each pair were identical in length, as were the embedded words contained within the targets. And third, for both targets within a given pair, the embedded word appeared in either the initial position (11 pairs: e.g., *army*) or the final position (7 pairs: e.g., *plot*).

Both target words from each pair were fitted into two different sentence frames. Therefore, each participant viewed all 36 target words. For each target word within a pair, a given sentence frame shared its beginning (up until the target word) but differed after the target word. Two different presentation lists were constructed such that if a target word appeared in sentence frame one on list one, it appeared in sentence frame two on list two, and vice-versa. For both presentation lists, 119 filler sentences were interspersed among the experimental sentences.

Additional norming studies were conducted to ensure that for each sentence frame, both target words fit equally well with the sentence beginning, but that neither target word was highly predictable. For the goodness of fit measure, 20 members from the University of Massachusetts community were provided with each target word along with one of its possible sentence beginnings (10 per sentence beginning). The participants were asked to rate on a 1 to 7 scale how well they felt that the word fit in the sentence beginning, with higher numbers signaling better fit. Sentences for which the target words did not receive similar ratings were revised and then re-normed (20 participants). For the final set of sentence frames, the mean goodness of fit was 5.66 for targets containing a higher frequency embedded word and 5.60 for targets containing a lower frequency embedded word ($p > .59$).

¹Due to the extensive amount of lexical embedding in the English language, it was impossible to exclude all words containing two-letter or one-letter function words. Thus, these were not controlled in the present experiment nor in the previous studies which have examined lexical embedding (e.g., Davis & Taft, 2005; Bowers et al., 2005).

For the predictability measure, 10 members from the University of Massachusetts community were presented with all of the original sentence beginnings and were asked to write a single noun that they felt could fit as the next word in the sentence. This process was repeated for the sentences which had been re-normed for the goodness of fit measure (nine participants). None of the participants provided the target word for any of the sentences in either of the predictability norming sessions (thus target predictability was zero).

Results

As is standard practice with eye movement experiments (Rayner, 1998), a number of dependent measures were examined. Probability of fixation reflects the percentage with which one or more first pass fixations occurred in the target region. First fixation duration represents a very early processing measure, the duration of the first fixation on the target word irrespective of the number of fixations. Gaze duration is the sum of all fixations on the target word prior to an eye movement away from the word. The remaining four measures reflect later stages of processing. Go-past time (also referred to as regression path duration) is the summed duration of all fixations from first entering the target word region to first going past it, including regressive fixations to earlier regions. Spillover is the duration of the first fixation following an eye movement away from the target word. Regressions back to the target word reflect the percentage with which the target word was re-fixated following a rightward eye movement, and second pass time reflects the amount of time spent re-reading the target word after it has been re-fixated.

Trials on which there was a track loss on the target word, or on the word that preceded or followed it (2.04% of the data), were not included in any of the analyses. The data were analyzed with the error term computed over participants (t_1) and items (t_2). Embedded word frequency was treated as a within-participants factor in the t_1 analyses. Because the two target words were well-matched on the characteristics in Table 1, embedded word frequency was treated as a within-items factor in the t_2 analyses. Prior to analyses, short fixations of less than 80 ms and long fixations of more than 1000 ms were removed by the data analysis software.

Table 2 contains the means for each of the dependent measures as a function of the frequency of the embedded word contained within the target. Although the probability of fixation on the target was similar for the two conditions (both p s > .31), fixation times were longer when the target word contained a higher frequency embedded word than when it contained a lower frequency embedded word. The 8 ms effect in first fixation was not significant: $t_1(29) = 1.84$, $SEM = 4.21$, $p = .076$; $t_2(35) = 1.19$, $SEM = 5.94$, $p = .24$. However, the 18 ms effect in gaze duration was fully significant: $t_1(29) = 2.71$, $SEM = 6.68$, $p = .011$; $t_2(35) = 2.10$, $SEM = 8.41$, $p = .043$. Turning to the later processing measures, there were no reliable effects in go-past time (both p s > .71), spillover (both p s > .45), percentage of regressions back to the target word (both p s > .17), or in second pass times (both p s > .28).

Discussion

The current results provide evidence that higher frequency embedded words interfere with target word processing during silent reading in English. This finding is consistent with the pattern obtained with the lexical decision task (Davis et al., 2009; Davis & Taft, 2005) and the semantic categorization task (Bowers et al., 2005; McCormick et al., 2010), and provides an important extension for the English language, demonstrating that embedded words have an interfering effect in English even when explicit judgments about lexicality or category membership are not solicited.

Given this important extension, we can conclude that embedded words have a similar, inhibitory, effect during silent reading in English and Spanish (Davis et al., 2009). Furthermore, we can now conclude that this effect generalizes to embedding positions other than the outer position used in Davis et al. (2009). In the current experiment the embedded words appeared in either the initial position (11 word pairs) or the final position (7 word pairs). It is interesting to note that although interference effects in lexical decision and semantic categorization have been observed for a range of positions (initial, middle, final, outer), these effects appear to be weaker when the embedded word appears in the final position (for a review see Davis et al., 2009). Post-hoc analyses were conducted to examine the effect of embedded word position on gaze durations. Although the interference effect observed in gaze durations was 25 ms for the initial pairs but only 5 ms for the final pairs, the embedded word frequency by position interaction was not significant, $F(1, 34) = 1.41, p > .24$. Given that only 7 word pairs contained an embedded word in the final position, and because the word pairs were not well-matched as a function of position (with respect to other characteristics known to influence word identification), clearly it is premature to generate conclusions about the relative influence of initial and final embeddings during silent reading.

Although this experiment was not designed to test any particular theory of reading, the results fit nicely within the conceptual framework posited by at least one model that explains the pattern of eye movements that occur during normal reading, the E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998). According to the E-Z Reader model, identification of word N during normal reading occurs in two stages, L1 and L2. In the context of a two-stage model of word processing (Paap, Newsome, McDonald, & Schvaneveldt, 1982), one might loosely associate L1 with the activation stage and L2 with the verification stage whereby competition processes among activated candidates occur until the correct candidate is selected (Davis et al., 2009; Williams et al., 2006). Furthermore, it is reasonable to assume that earlier eye movement measures are primarily a reflection of processes that underlie the activation stage and that later measures are associated with the verification stage. Given this proposal, the current finding that embedded word frequency had an effect on gaze durations but not first fixation durations suggests that lexical embeddings influence processes mainly associated with the verification stage and that selecting the correct candidate takes more time when a higher frequency embedded word is present (see Davis et al., 2009 for a similar suggestion).

Although the current results clearly demonstrate that lexical embeddings influence word identification during silent reading, they alone do not shed light on what characteristics of the embedded words are responsible for the inhibitory effects that have been observed. A contrast of the present results with those obtained from eye tracking investigations of embedded words that are morphologically related to the whole word is instructive. In these studies (e.g. Juhasz et al., 2003; Niswander-Klement & Pollatsek, 2006; Niswander et al., 2000) fixation durations on morphologically complex words were shorter when they contained a high frequency root or lexeme (e.g., *blindness* vs. *bleakness*). However, this morphological effect does not necessitate a semantic relationship, as studies have reported facilitation for high frequency lexemes in semantically opaque compound words (e.g. *deadline*), when the lexemes do not overlap in meaning with the entire compound expression (Juhasz, 2007; Pollatsek & Hyönä, 2005). Thus, it appears to be the lack of a morphological relationship that is crucial for interference to be observed from a high frequency embedded word. Additional research is necessary to understand at what point in the word recognition process the morphological status of the currently fixated word begins to modulate the influence of embedded words. Since high frequency morphemes have been found to shorten the very first fixation in morphologically complex words (e.g. Juhasz et al., 2003; Juhasz, 2007, Niswander et al., 2000), the morphological status of the currently fixated word must

constrain word recognition processes quickly – perhaps through a process of rapid morpho-orthographic decomposition applied to letter sequences that have an apparent morphological structure (Rastle & Davis, 2008).

A contrast of the present results with those obtained from recent eye tracking investigations of embedded words that are morphologically related to the whole word is instructive. In these studies fixation durations on morphologically complex words were shorter when they contained a high frequency root or lexeme (e.g., Juhasz, 2007). Thus, it appears that a high frequency embedded word interferes with word identification when it is not morphologically related to the whole word, but that it facilitates word identification when a morphological relationship does exist.

Continuing with the topic of what characteristics of an embedded word might influence word identification, future research should examine the relationship between the embedded word's phonological representation and that of the whole word. In lexical decision and semantic categorization studies (Bowers et al., 2005; Davis & Taft, 2005), a high frequency embedded word has been shown to produce interference regardless of whether its pronunciation was similar or dissimilar to that of the whole word (e.g., *treaty* vs. *cello*). Thus, to the extent that an embedded word's phonological representation is activated, it appears that the degree of similarity in pronunciation between it and the whole word might not be critical in producing the interference effects. However, it remains to be determined whether this conclusion generalizes to silent reading.

We conclude with a point regarding the general implications these results have for the conceptualization of orthographic similarity. Orthographic neighbors have traditionally been defined as words that differ from the base word by the substitution of a single letter. However, a number of findings converge on the idea that orthographic similarity exists beyond letter substitutions (e.g., Bowers et al., 2005; Davis et al., 2009; Davis & Taft, 2005; McCormick et al., 2010; see also work on transposed letter effects, e.g., Johnson, Perea, & Rayner, 2007). The current findings add to this growing body of evidence, suggesting that a word's orthographic neighborhood includes words embedded within itself.

Acknowledgments

This research was supported by Grant HD26765 from the National Institute of Health. Barbara J. Juhasz and Kristin M. Weingartner were supported by National Institute of Mental Health Training Grant MH16745 awarded to the University of Massachusetts, Amherst. These data were collected when all authors were at the University of Massachusetts, Amherst. We thank Jeff Bowers, Manuel Perea, and an anonymous reviewer for comments on an earlier draft.

References

- Andrews S. The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*. 1997; 4:439–461.
- Balota DA, Yap MJ, Cortese MJ, Hutchison KA, Kessler B, Loftis B, Neely JH, Nelson DL, Simpson GB, Treiman R. The English Lexicon Project. *Behavior Research Methods*. 2007; 39:445–459. [PubMed: 17958156]
- Balota, DA.; Cortese, MJ.; Hutchison, KA.; Neely, JH.; Nelson, D.; Simpson, GB.; Treiman, R. The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords. Washington University; 2002. <http://elexicon.wustl.edu/>
- Bowers JS, Davis CJ, Hanley DA. Automatic semantic activation of embedded words: Is there a “hat” in “that”? *Journal of Memory & Language*. 2005; 52:131–143.
- Coltheart, M.; Davelaar, E.; Jonasson, JT.; Besner, D. Access to the internal lexicon. In: Dornic, S., editor. *Attention and performance*, VI. Hillsdale NJ: Erlbaum; 1977. p. 535-555.

- Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J. DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*. 2001; 108:204–256. [PubMed: 11212628]
- Davis CJ, Perea M, Acha J. Re(de)fining the orthographic neighborhood: The role of addition and deletion neighbors in lexical decision and reading. *Journal of Experimental Psychology: Human Perception and Performance*. 2009; 35:1550–1570. [PubMed: 19803656]
- Davis CJ, Taft M. More words in the neighborhood: Interference in lexical decision due to deletion neighbors. *Psychonomic Bulletin & Review*. 2005; 12:904–910. [PubMed: 16524009]
- Gilhooly KJ, Logie RH. Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods and Instrumentation*. 1980; 12:395–427.
- Harm MW, Seidenberg MS. Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*. 2004; 111:662–720. [PubMed: 15250780]
- Johnson RL, Perea M, Rayner K. Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance*. 2007; 33:209–229. [PubMed: 17311489]
- Juhasz, BJ. The influence of semantic transparency on eye movements during English compound word recognition. In: vonGompel, R.; Murray, W.; Fischer, M., editors. *Eye movements: A window on mind and brain*. Boston, MA: Elsevier; 2007. p. 373–389.
- Juhasz BJ, Starr M, Inhoff AW, Placke L. The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions, and eye fixations. *British Journal of Psychology*. 2003; 94:223–244. [PubMed: 12803817]
- Kucera, H.; Francis, W. *Computational analysis of present-day American English*. Providence, RI: Brown University Press; 1967.
- McCormick SF, Davis CJ, Brysbaert M. Embedded words in visual word recognition: Does the left hemisphere see the rain in brain? *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2010; 36:1256–1266.
- Nation K, Cocksey J. Beginning readers activate semantics from sub-word orthography. *Cognition*. 2009; 110:273–278. [PubMed: 19108820]
- Niswander-Klement E, Pollatsek A. The effects of root frequency, word frequency, and length on the processing of prefixed English words during reading. *Memory & Cognition*. 2006; 34:685–702.
- Niswander E, Pollatsek A, Rayner K. The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes*. 2000; 15:389–420.
- Paap KR, Newsome SL, McDonald JE, Schvaneveldt RW. An activation verification model for letter and word recognition: The word-superiority effect. *Psychological Review*. 1982; 89:573–594. [PubMed: 7178333]
- Paterson KB, Livensedge SP, Davis CJ. Inhibitory neighbor priming effects in eye movements during reading. *Psychonomic Bulletin & Review*. 2009; 16:43–50. [PubMed: 19145008]
- Perea M, Acha J, Fraga I. Lexical competition is enhanced in the left hemisphere: Evidence from different types of orthographic neighbors. *Brain and Language*. 2008; 105:199–210. [PubMed: 17905425]
- Perea M, Pollatsek A. The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*. 1998; 24:767–779. [PubMed: 9627415]
- Pollatsek A, Hyönä J. The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes*. 2005; 20:261–290.
- Pollatsek A, Perea M, Binder K. The effects of neighborhood size in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*. 1999; 25:1142–1158. [PubMed: 10464947]
- Rastle K, Davis MH. Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*. 2008; 23:942–971.
- Rayner K. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*. 1998; 124:372–422. [PubMed: 9849112]

- Reichle ED, Pollatsek A, Fisher DL, Rayner K. Towards a model of eye movement control in reading. *Psychological Review*. 1998; 105:125–157. [PubMed: 9450374]
- Sears CR, Campbell CR, Lupker SJ. Is there a neighborhood frequency effect in English? Evidence from reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*. 2006; 32:1040–1062. [PubMed: 16846296]
- Slattery TJ. Word misperception, the neighborhood frequency effect, and the role of sentence context: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*. 2009; 35:1969–1975. [PubMed: 19968447]
- Williams CC, Perea M, Pollatsek A, Rayner K. Previewing the neighborhood: The role of orthographic neighbors as parafoveal previews in reading. *Journal of Experimental Psychology: Human Perception and Performance*. 2006; 32:1072–1082. [PubMed: 16846298]

Appendix

Note: Target words appear in italics. For each sentence frame, the version containing the target with a higher frequency embedded word appears first, followed by the version containing the target with a lower frequency embedded word.

1. Rachel saw that the *barn* was a mess and needed to be cleaned.
 Rachel saw that the *bunk* had already been made up by her son.
 Johnny collapsed in the *barn* from a massive heart attack.
 Johnny collapsed in the *bunk* from extreme exhaustion.
2. Stacy's mom told her about the *plot* that she had been writing for her book.
 Stacy's mom told her about the *tale* that she had learned during her childhood.
 Margaret disliked the *plot* and thought it was too predictable.
 Margaret disliked the *tale* and requested that it be rewritten.
3. Karen slowly drew a *card* from the magician's hands.
 Karen slowly drew a *king* from the deck of cards.
 Tom hoped for a *card* but everyone forgot his birthday.
 Tom hoped for a *king* but he only got a jack instead.
4. Tim's father left the *army* after he hurt his back in an accident.
 Tim's father left the *firm* after he found a better job.
 Rick did not like being a part of the *army* and really wanted to quit soon.
 Rick did not like being a part of the *firm* and thought about changing careers.
5. Nancy spotted the *flaw* in the suit she tried on when she was shopping.
 Nancy spotted the *mink* with a very high price when she was shopping.
 The photographer thought that the *flaw* in the picture was horrible.
 The photographer thought that the *mink* in the picture was lovely.
6. Oscar could not help but stare at the ugly *scar* that the man had on his arm.
 Oscar could not help but stare at the ugly *lamp* that the man had on his desk.
 Jeff thought that the *scar* above his eyebrow made him look distinguished.
 Jeff thought that the *lamp* above his desktop computer was too bright.

7. The farmer searched for a type of *wheat* that would survive without much water.
The farmer searched for a type of *breed* that would survive the cold winter.
I was shown a new type of *wheat* that does not need as much sunlight.
I was shown a new type of *breed* that combines a poodle and a beagle.
8. Jim saw that the *broom* was not in the closet where it was supposed to be.
Jim saw that the *crust* was still on the bread and refused to eat the sandwich.
Melinda looked at the *broom* on her porch and realized it was broken.
Melinda looked at the *crust* on her sandwich and realized it was moldy.
9. The owner of the *manor* was very pleased with her new purchase.
The owner of the *kitty* was very pleased with her new furry friend.
Christopher heard that the *manor* had been purchased last week.
Christopher heard that the *kitty* had been rescued from a tree.
10. Tom heard about the *carol* Shelly sang at the local mall.
Tom heard about the *binge* Shelly went on last week.
Martha told us about the *carol* she sang to the elderly at Christmas.
Martha told us about the *binge* she went on before she became sober.
11. Jeremy took the *sword* from the wall as he prepared to go to battle.
Jeremy took the *cream* from the refrigerator to use in his coffee.
Arthur picked up the *sword* and defended his castle from attack.
Arthur picked up the *cream* and poured it into his freshly brewed coffee.
12. Bob realized that the *artery* was very blocked and needed to be treated.
Bob realized that the *bundle* was very large and decided to split it in half.
Lisa knew that the *artery* in her dad's forehead was throbbing out of anger.
Lisa knew that the *bundle* in her dad's pocket was actually stolen money.
13. Tim read about the widely publicized *artist* in the entertainment magazine.
Tim read about the widely publicized *matter* in the local newspaper.
Maggie asked her secretary if the important *artist* had been paid yet.
Maggie asked her secretary if the important *matter* had been resolved yet.
14. The teacher used an odd *manner* of speech which made the students laugh.
The teacher used an odd *figure* of speech to get his point across.
Nancy did not appreciate the *manner* that her daughter was using.
Nancy did not appreciate the *figure* that was used in the textbook.
15. Yesterday Tina's *agency* took a beating after the stock had taken a dive.
Yesterday Tina's *finger* took a beating after she slammed the car door on it.
I was told that Kathy's *agency* was recently put out of business.

I was told that Kathy's *finger* was broken in the accident.

16. Michael bought a small *rocket* for his nephew to assemble.

Michael bought a small *racket* for his nephew to play tennis with.

The fire destroyed the expensive *rocket* owned by the American government.

The fire destroyed the expensive *racket* owned by the famous tennis player.

17. Janice's teacher described the *syntax* that provided the rules of grammar.

Janice's teacher described the *coffin* that King Tut had been buried in.

Ken thought that the *syntax* in the computer program was difficult to read.

Ken thought that the *coffin* in the horror film was very ornate and beautiful.

18. The apprentice went to the *cellar* to get some wine for dinner.

The apprentice went to the *master* to get some advice about tools.

Michelle was told that the *cellar* of the house was very large.

Michelle was told that the *master* of the house was very nice.

Table 1

Characteristics of the Target Words

	Higher-frequency embedded word	Lower-frequency embedded word
Length	5.06 (.87)	5.06 (.87)
Frequency (per million) ^a	32.44 (39.31)	53.67 (82.41)
Neighborhood size (N)	4.11 (4.23)	6.22 (5.09)
Number of higher-frequency neighbors	1.64 (2.37)	2.69 (3.55)
Familiarity	6.43 (.71)	6.65 (.26)
Concreteness	5.14 (1.79)	5.23 (1.68)
Age of acquisition	3.67 (1.29)	3.39 (1.33)

Note: Values are means, with standard deviations in parentheses.

^aTaken from Ku era and Francis (1967)

Table 2

Eye Movement Measures for the Target Words

	Higher-frequency embedded word	Lower-frequency embedded word
Probability of fixation (%)	87.2 (11.6)	85.0 (11.8)
First fixation	252 (28)	244 (25)
Gaze	276 (44)	258 (27)
Go-past time	301 (61)	297 (65)
Spillover	248 (34)	245 (20)
Regressions back to target word (%)	11.5 (9.8)	9.3 (9.3)
Second pass	34.6 (33.9)	29.4 (31.0)

Notes: Standard deviations appear in parentheses. Mean duration times are presented in milliseconds.