The effects of root frequency, word frequency, and length on the processing of prefixed English words during reading

ELIZABETH NISWANDER-KLEMENT and ALEXANDER POLLATSEK
University of Massachusetts, *Amherst, Massachusetts*

Two experiments were performed to investigate the roles of root morpheme frequency and word frequency in the encoding of prefixed words in sentence context. Two alternative words, which were equated on other indices but differed with respect to either root morpheme frequency or word frequency, were embedded in the same sentence frame. In Experiment 1, there was a significant root morpheme frequency effect on gaze duration but only a marginal word frequency effect. Post hoc analyses indicated that word length influenced both effects, with word frequency effects predominating for shorter words and root morpheme effects predominating for longer words. In Experiment 2, word length was also manipulated. There was a significant root frequency effect for longer prefixed words and a significant word frequency effect for shorter prefixed words. The results were best explained by a dual-route model (with competing whole-word access and compositional routes), in which increasing word length inhibited the whole-word process, relative to the compositional route.

Although most of the research on word identification has been on monomorphemic words, there has been an increasing number of studies in which the processing of multimorphemic words has been explored since the pioneering work of Taft (1979) and Taft and Forster (1975, 1976). The major issue in this research has been the role that component morphemes play in the identification of multimorphemic words. There are a variety of models that have been proposed for the access and storage of complex words in the mental lexicon. Most fit into three basic categories: (1) direct access, (2) decomposition, and (3) dual-route (or dual-access) models involving both full-form storage and some form of decomposition within a dual-process system. (Models with a parallel distributed processing architecture, such as that in Seidenberg & McClelland, 1989, are somewhat harder to characterize but probably are most similar to the dual-route models, in that they typically posit that morpheme-like entities are involved in parallel with letter entities in word recognition.) In the direct access model (e.g., Butterworth, 1983; Giraudo & Grainger, 2000), a complex word has an individual full-form representation stored in the lexicon,

and it is this full-form representation that is involved in initial access (i.e., a complex word is accessed no differently than a simple, monomorphemic word). These models posit that morphological components are activated only after initial access but can influence postlexical processing.

Decompositional models, in contrast, posit that a morphemically complex word is encoded by a process in which the whole word is necessarily decomposed into its parts. The purpose of the decomposition is the extraction of the word's morphemes, and these models often include multiple levels (e.g., word and morpheme) within the lexicon. There are two basic types of decompositional models. One type, which might be termed a *fully decompositional* model, is one in which each morphemic component is accessed and then the full-form representation is constructed from them. Clearly, such a mechanism is almost necessary to understand a novel complex word, such as *mouseball*; however, it seems unlikely as an explanation for understanding the meaning of any complex words in the lexicon, with the possible exception of words that are completely orthographically and semantically transparent, such as *uncover*. There are few, if any, compound words that are completely transparent. For example, *cowboy*, which is often given as an example of a transparent word in English, could be a synonym for *calf*. That is, it is only transparent in the sense that the meaning of *cowboy* is related to cows and to males. One possible way to maintain a fully decompositional model is to posit that the meaning is not necessarily constructed from the parts. Such a model is Taft's (2004) model. The model has three levels of nodes. The first contains the form code, the word's orthography and phonology. The second is the lemma level, and a distinction is made at this level between the activation of trans-

parent and opaque affixed words. A completely transparent affixed word, such as *reheat*, does not have its own representation at the lemma level and, instead, activates the lemmas *re* and *heat*. In contrast, a word that is not completely transparent and cannot be predicted from its constituents, such as *feathery*, is represented at the lemma level, and the meaning is accessed from there.

A second type of decompositional model is Taft and Forster's (1976) *file drawer* model, in which a full-form representation for each complex word is stored in the lexicon but is accessed by first extracting and accessing the word's root morpheme. When a complex word is encountered, accessing the root morpheme is likened to accessing a file drawer. This is followed by a serial search through the drawer for the whole complex word's representation, in which the whole-word representations are filed in descending order of frequency and the search follows this order. A variant of the file drawer model is Lukatela, Gligorijević, Kostić, and Turvey's (1980) satellite entries model (see also Caramazza, Laudanna, & Romani, 1988). (Although this model was specifically designed for Serbo-Croatian, in which all complex word forms were hypothesized to be accessed through their singular noun form, the general framework can be generalized across words and languages by assigning the nucleus role to the root morpheme.) The first step of processing is similar to that of the file-drawer model, except that the file drawer is now a nucleus around which all of its forms' representations are clustered as satellites. In contrast to the serial processing proposed in the file drawer model, the satellites are accessed in parallel. Instead of each entry being individually sorted through in order of decreasing frequency, the satellites have activation thresholds that are negatively correlated to their relative frequencies. The model's predictions are quite similar to those of the file drawer model.

The third category of complex word encoding models is *dual-route* (or dual-access) models (Hyönä & Pollatsek, 1998; Pollatsek, Hyönä, & Bertram, 2000), in which there are two routes of access to the lexicon for complex words. One is a direct access route for which the whole word is the access unit used to directly access the word's full-form representation; the other is a decompositional route, and this route, as our discussion of decompositional models above suggests, can take many forms. Moreover, in such a dual-route framework, it is no longer implausible that the decompositional process is one that has no full-form representation. How these *dual routes* operate has been a topic of debate. One issue is whether or not they operate in parallel (the and/and model) or whether one route is selected and the alternate route is taken only if the first route fails (the either/or model). This raises additional questions. If the two routes operate serially, as in the either/or model, how is it determined which route to activate first (Lukatela et al., 1980)? Finally, if both routes operate in parallel, an unanswered question still open to debate is whether or not these routes operate independently (Dominguez, Cuetos, & Segui, 1999) or have an interactive relationship (Deutsch, Frost, & Forster, 1998; Pollatsek et al., 2000).

A major source of data that has the potential for discriminating among these models is whether the time to identify a complex word is (1) influenced by the frequency of the constituent morphemes, (2) influenced by the frequency of the whole word, or (3) influenced by both frequencies. In addition, the relative time course of these influences is potentially valuable data for discriminating among these models. Since the experiments in this study examine the processing of morphemically complex words, using eye movement measures while people read sentences containing a morphemically complex target word, we will largely confine our discussion to these experiments. We think that this is the most ecologically valid way of assessing these theories. Moreover, the sequence of eye fixations allows one to examine the time course of processing in a nonintrusive way (see Rayner, 1998). That is, many morphemically complex words are fixated more than once, so that one can examine whether a particular type of frequency effect, for example, is seen on the first fixation or whether it surfaces only on later fixations. That is not to say that we view individual fixations as processing stages; there is no reason why several processing stages could not be completed on the same fixation.

The basic paradigm for testing whether the frequency of a component morpheme influences word processing is to choose two words that differ in the frequency of a component morpheme (holding the frequency of the word, the length, and other relevant parameters constant) and embed them in the same sentence context (Hyönä & Pollatsek, 1998; Pollatsek et al., 2000). Many measures can be used to assess the word identification process, but the most relevant are the *first-fixation duration* on the target word, the *gaze duration* (the sum of all fixations on the target before the word is left in either direction), and the *total time* (the total fixation time on the target word, including revisits to the target word after fixations elsewhere). A similar paradigm is used for assessing word frequency effects when the words in a pair differ on word frequency, with other parameters controlled (Alegre & Gordon, 1999; Niswander, Pollatsek, & Rayner, 2000; Pollatsek et al., 2000). This is the methodology in all the studies discussed below, unless an exception is noted. Each participant sees each sentence frame only once and, thus, sees only half the target words in each condition.

Most of the experiments using eye movements have employed compound words. As a result, we will discuss these first. The initial studies employed Finnish two-constituent compound words, which are always written without a space between the constituents (Hyönä & Pollatsek, 1998; Pollatsek et al., 2000). In three separate experiments, the following were independently varied: (1) the frequency of the first constituent, (2) the frequency of the second constituent, and (3) the frequency of the whole word. (The constituents are actually often not a single morpheme, since they have inflectional endings.) All three frequencies—the first-constituent, second-constituent, and whole-word frequency—produced large effects on the gaze duration on the compound words, with gaze durations in the lower

frequency conditions that were close to 100 msec longer than those in the higher frequency conditions. However, the effects followed a somewhat different time course. The effect of the first-constituent frequency occurred early (on first-fixation duration), but the effects of the second-constituent frequency and whole-word frequency surfaced somewhat later in the eye movement record. Thus, on the face of it, both componential processes and direct access were involved, with componential processes having an earlier influence. We will discuss the implications more fully below.

All of the studies above employed long words (12 letters or more), and it may be that componential processing is enhanced when all of the letters cannot be easily perceived on a fixation. Bertram and Hyönä (2003) tested this hypothesis by experimentally manipulating length factorially (8 or fewer letters vs. 12 or more letters) with either the first-constituent frequency or the whole-word frequency of compound Finnish words. The data indicated that constituent frequency had a larger effect for the long compound words than for the short compound words. In contrast, word frequency had a larger effect on the short compound words than on the long compound words. This suggests that longer Finnish compound words are more likely to be processed using their constituent morphemes, whereas shorter Finnish compound words are more likely to be processed by their full forms. This result is consistent with a dual-process model but is less consistent with either of the decomposition models, since the order of processing is not posited to be influenced by word length.

There have been two studies in English of compound English words written without a space between constituents (e.g., *battlefield*, *blackjack*), in which essentially the same paradigm was used. Juhasz, Starr, Inhoff, and Placke (2003) investigated the effect of constituent frequencies and found a robust 34-msec effect of second-constituent frequency on gaze duration but only a nonsignificant 11-msec effect of first-constituent frequency on first-fixation duration. The robust second-constituent and elusive first-constituent frequency effects were consistent with the data from their lexical decision and naming experiments. Andrews, Miller, and Rayner (2004) also manipulated the constituents' frequencies and found significant effects of both first-constituent frequency and second-constituent frequency on gaze duration and total time and a marginal effect of first-constituent frequency on first-fixation duration. Regression analyses also indicated a contribution of whole-word frequency. These studies suggest that the processing of compound words in reading is similar in Finnish and English and that the chief difference is that Finnish compound words are generally longer than English (unspaced) compound words.

These findings clearly rule out a direct access model as an explanation for how compound words are processed, because it predicts that morphemic frequency effects should occur only after whole-word frequency effects. As a result, we will not discuss direct access models any further. The data are at least qualitatively compatible with a dual-route model in which there is a *race* between a *direct look-up* process and a decompositional process, where the decompositional process gets started early in the race. The first-constituent frequency effect is compatible with a file drawer model if one assumes that access of the first constituent is access of the file drawer, and the word frequency effect is also compatible with a file drawer model if one assumes that the order of words in the file drawer is likely to be confounded with word frequency. It is less clear whether such a model predicts second-constituent frequency effects, however, since the frequency of the second constituent (with whole-word frequency held constant) would not obviously be related to how high up in the file drawer the word was. Most of these effects are roughly compatible with Taft's (2004) full decomposition model. However, the fact that whole-word frequency influences processing at least as early as second-constituent frequency does seems to present a problem, since whole-word frequency in the model should come into play only after both constituents have been accessed.

Word and root frequencies were also independently manipulated in order to study the processing of inflected and derived suffixed words in English (Niswander et al., 2000). For the derived suffixed words, there was an effect of root frequency on first-fixation duration, and word frequency had a large effect on gaze duration. This pattern of results is similar to that in the Finnish data on compound words, in the sense that the root morpheme frequency effect appeared early in processing (on first-fixation duration) and the whole-word frequency effect appeared later. There was a somewhat different and unexpected set of results for the inflected words. Overall, high (word) frequency inflected words had shorter gaze durations than did low-frequency inflected words. The pattern was more complicated for root frequency. As was expected, there was an overall effect of root frequency on first-fixation duration, but when the nouns and the verbs were analyzed separately, the pattern was not consistent. It appeared that root frequency affected reading times for the nouns, but not for the verbs. However, further analysis of the materials indicated that some of the verbs contained roots that were used more often as nouns than as verbs, and the occurrence of a conflict between which version to access (noun or verb) was proposed to explain the findings. Word length may again play a role, since the effects for these inflected and derived words were not that large. Another relevant variable may be the length of the morpheme being studied relative to the length of the whole word. In the case of English inflected words, that difference is one or two characters, which may make processing of the whole word as a unit almost as easy as processing the root morpheme.

No study of frequency effects with prefixed words has been published in which the participants' eye movements were monitored during sentence reading. However, Beauvillain (1996) studied the effects of root and word frequency with prefixed and suffixed words by monitoring participants' eye movements during a semantic relatedness task in which words were used in isolation. Word frequency effects on gaze duration were found for both prefixed and suffixed words, and a root frequency effect

was found for suffixed words on first-fixation duration. In contrast, a root frequency effect for prefixed words occurred later—on second-fixation duration. The time course of the root frequency effects is consistent with the placement of the root within the word and with the findings for suffixed and compound words presented so far.

Summary

The previous research provides evidence of both root and word frequency effects with morphemically complex words. Furthermore, the time course of the root (or constituent) frequency effects indicated that the frequency of either the first constituent of a compound word or the root of a suffixed derived word appeared to have a somewhat earlier effect than did the frequency of the whole word. These findings are inconsistent with a direct access model and are at least roughly consistent with either a dual-route model or a file drawer type of compositional model of complex word processing. They seem somewhat less consistent with a fully decompositional model. In addition, for compound words, word length appears to modulate the relative effects of first-constituent and whole-word frequency, with the former effects emerging strongly only for longer compounds. This seems compatible with a dual-route model but, on the surface, less compatible with a file drawer model, in which access of the initial constituent appears to be an obligatory first step in encoding. However, the processing of compound words may be different from the processing of other morphemically complex words.

The present experiments represent an attempt to broaden our knowledge of how morphemically complex words are encoded in the process of normal reading by examining the processing of prefixed words. In Experiment 1, there were two manipulations: (1) The root frequency of the prefixed word was manipulated while the whole-word frequency was held constant, and (2) the whole-word frequency was manipulated while the root frequency was held constant. Experiment 2 also systematically varied the length of the prefixed words, in order to examine whether word length influences the relative contribution of whole-word and compositional processes in encoding in a way similar to the processing of Finnish compound words.

EXPERIMENT 1

As has been indicated above, the basic logic of the method is as follows. Two words that are otherwise equated for stimulus characteristics but vary on either root morpheme frequency or word frequency are inserted in the same sentence frame. If one obtains a difference in processing time between two prefixed words that differ in the frequency of their roots but are otherwise matched (including being equated on word frequency and having the same prefix), it indicates that accessing the root morpheme plays a role in the identification of the word. If one obtains an analogous word frequency effect (when the root frequency and the prefixes are matched), it likely indicates that there is some role for the whole word in the access process and that processing is not completely compositional (i.e., looking up the root and then pasting on the prefix).

Method

Participants. Thirty-six native English speakers from the University of Massachusetts community participated in the experiment. Each participant received either course credit or money for his or her participation.

Apparatus. Eye movements were recorded by a Fourward Technologies Dual Purkinje Eyetracker, which has a resolution of less than 10 min of arc and an output that is linear over the angle subtended by a line of text. The eyetracker was interfaced with a Pentium computer. The computer was also interfaced with a ViewSonic 17G display on which the sentences were presented. The display was 61 cm from the participant's eye, and four characters equaled 1° of visual angle. Viewing was binocular, but eye movements were recorded from the participant's right eye. A bitebar was used to eliminate head movements in the experiment.

Materials. Prefixed English words were used as target words. The words ranged in length from 5 to 14 characters and consisted of a prefix plus a free root morpheme. A free root morpheme is a root that is able to stand alone as a word (e.g., *move* in *remove*), in contrast to a bound root morpheme (e.g., *ceive* in *receive*). In all the target words, the full root word was transparently represented in the orthography, so that no letters were added or deleted for the attachment to the prefix. Finally, an effort was made to include a range of prefixes; therefore, words containing the following prefixes were included: *de-, un-, re-, dis-, in-,* and *im-*.

There were two sets of word pairs: one in which root frequency was manipulated with word frequency controlled, and a second set in which word frequency was manipulated with root frequency controlled. The stimulus characteristics are displayed in Table 1. The prefixed word and root word frequencies were determined using Francis and Kučera (1982). The root frequency was defined as the frequency of the root word alone (e.g., the frequency of the word *move* as the root frequency for *remove*), and the word frequency was defined as the frequency of the whole word. (Since other researchers have defined root frequency differently, we will present analyses later using other definitions.) The two words in each pair shared the same prefix and either were equal in length or differed in length by one character. (The complete set of stimuli are given in Appendix A.) As can be seen in Table 1, root frequency differed substan-

Table 1
Experiment 1: Stimulus Characteristics

Stimulus Characteristic	Root Frequency				Word Frequency			
	Low		High		Low		High	
	M	SD	M	SD	M	SD	M	SD
Mean length	8.1	2.14	7.9	2.46	7.6	1.50	7.6	1.54
Mean word frequency	2.4	1.86	3.4	2.06	4.3	5.54	46.0	102.20
Mean root frequency	9.1	7.94	199.6	399.10	146.0	109.40	303.6	358.40

tially between the two words in each pair in the root frequency set, whereas mean whole-word frequency was approximately equated. Conversely, in the word frequency set, whole-word frequency differed substantially between the two words in each pair, whereas the mean root frequency was approximately equated.[1]

Each pair of target words was embedded in a sentence frame, with a minimum of two words before and after a target word in every sentence. Because the two target words within each pair were placed alternately within the same sentence frame, there were two lists of sentences constructed, so that each participant saw only one word from each pair and saw each sentence frame only once. In addition, the frequency conditions were counterbalanced across both lists, so that each participant saw items from all four conditions. This resulted in each participant's seeing 36 experimental sentences: 10 in each of the two word frequency conditions and 8 in each of the root frequency conditions. In addition, there were 84 filler sentences, and the order of the 120 sentences seen (36 experimental and 84 filler) was randomized separately for each participant.

Due to a concern over the presence of similar themes in some of the sentence frames, a blocked design was used. The sentences were divided into two blocks, so that sentences that were considered to be similar were separated between the two blocks; then the remaining sentences were evenly divided between the two blocks. The blocks were constructed so that each block contained an equal number of sentences from each condition and an equal number of filler sentences. The order in which the blocks were presented was varied, so that half the participants saw Block Order 1 (Blocks A–B), whereas the other half saw Block Order 2 (Blocks B–A). The counterbalancing lists were also divided across the blocks and block orders. The result was sets of stimuli presented equally across participants: (1) Block Order 1, Counterbalancing List 1, (2) Block Order 1, Counterbalancing List 2, (3) Block Order 2, Counterbalancing List 1, and (4) Block Order 2, Counterbalancing List 2.

To make sure that the differences between the conditions were not caused by a difference in how well the two target words fit into their sentence frames, the experimental sentences were rated by a separate set of participants. Twenty-one participants were asked to read the sentence pairs and rate the relative naturalness of the sentences. More specifically, they were asked to respond that (1) the sentence containing the low-frequency word was more natural, (2) the sentence containing the high-frequency word was more natural, or (3) the sentences were equally natural. Each response was numerically coded as −1, +1, or 0, respectively. The order of the sentences was balanced, so that for half the items, the sentence containing the low-frequency target was listed first and, for the other half of the items, the sentence containing the high-frequency target was listed first. Moreover, the order was counterbalanced, so that for each item, half the participants saw the low-frequency target first and half saw the high-frequency target first. The matching was very good for both sets, since the means of the naturalness difference scores for the root frequency set and the word frequency set were +0.09 and −0.02, respectively.

Procedure. When the participants arrived for the experiment, they read a description of the procedure. A bitebar was then prepared, and the procedure was explained verbally in more detail. The eyetracker was then calibrated (which takes about 5 min). They were told that they would be given sentences to read and were asked to read each sentence for normal comprehension. To ensure comprehension, they were asked to answer a question about the sentence they had just read after 20% of the sentences.

Results

Data loss and exclusion criteria. Items for which track losses occurred around the target word were excluded from the analyses. To take care of the possibility that this would lead to cases in which excessive data loss occurred, a data loss criterion of 5% was set, and 6 participants were excluded on this basis. In addition, individual

fixations that had durations less than 80 msec or more than 800 msec were excluded from the final data set.

The purpose of collecting the comprehension question data was to assess whether the participants were following the instructions and reading the sentences. A criterion was set so that any participant who answered fewer than 75% of the comprehension questions correctly would be excluded from the analyses. On average, the participants scored over 97% on the comprehension questions, and no participant was excluded on this basis.

Eye movement measures. The Results sections will be organized in temporal order. We will begin with the earliest measure of processing, the probability of skipping the target word, followed by the first-pass measures (measures made during the first pass that the eyes make on a target word). Two first-pass measures are *first-fixation duration* (the duration of the first fixation on the target word) and *gaze duration* (the summed duration of all fixations on the target word before it is left in either direction). These are followed by two slightly later measures of target processing: *go-past* (which counts all fixations between the time at which the target word is first fixated on the first pass and the time at which a saccade leaves it to the right, including regressions back from the target word) and *spillover duration* (the duration of the first fixation immediately following the first pass on the target word). We will finish with two measures that include post–first-pass fixations on the target word; *second-pass time* (the mean duration of all fixations on the target word after the reader has left the target word to the right[2]) and *total time* (the summed duration of all fixations on the target word). We first will examine the root frequency set and then the word frequency set.

Root frequency set. The root frequency data for Experiment 1 are presented in Table 2. Although the probability that the target word was skipped was quite low in both the high and the low root frequency conditions, there was a significant difference [$t_1(35) = 2.03, p < .05$; $t_2(15) = 2.47, p < .05$]. (Since this difference in skipping rates did not replicate in Experiment 2, we will not comment on it further.) There was also a small (7-msec) effect of root frequency on first-fixation duration; however, it was not close to significant (ts < 1). In contrast, there was a large, significant root frequency effect on gaze duration. Words with low-frequency roots were fixated 54 msec

Table 2
Experiment 1: Reading Measures as a Function of Root Frequency

Reading Measure	Root Frequency	
	High	Low
Probability of skipping target word (%)	1.0	3.9
First-fixation duration (msec)	286	292
Gaze duration (msec)	357	411
Go-past time (msec)	407	457
Spillover duration (msec)	264	272
Second-pass time (msec)	69	74
Total time (msec)	427	486
Initial landing position on target word	3.33	3.16
Time of last fixation prior to fixating target (msec)	242	236

longer than words with high-frequency roots [$t_1(35) = 3.64, p < .01; t_2(15) = 2.53, p < .05$]. A somewhat later measure of processing is go-past. By this measure, there was a 50-msec root frequency effect [$t_1(35) = 2.50, p < .025; t_2(15) = 2.05, p < .10$]. Thus, root frequency had a large and significant effect on the first-pass reading times for prefixed words. There was little effect of root frequency once the readers had left the target word region, even as early as the next fixation. The spillover duration in the low- and high-frequency conditions differed only by 8 msec ($ts \leq 1$), and the second-pass time differed only by 5 msec ($ts < 1$). Finally, there was a sizable root frequency effect on total time. Low root frequency words were fixated a total of 59 msec longer than high root frequency words [$t_1(35) = 3.31, p < .01; t_2(15) = 2.24, p < .05$].

We also examined two other indices of word processing; the landing position on the target word and the duration of the fixation prior to the landing on the target word. For landing position, there were virtually no differences between root frequency conditions [$t_1(35) = 1.03, p > .20; t_2 < 1$]. Similarly, for the prior fixation duration, there was virtually no difference [$t_1(35) = 1.12, p > .20; t_2 < 1$].

Word frequency set. As with the root frequency manipulation, word frequency had little effect on the earliest measures of processing (see Table 3). The target words were rarely skipped, and there was virtually no difference between the high- and the low-frequency words ($ts < 1$); the first-fixation duration for the low-frequency words was only 6 msec more than that for the high-frequency words ($ts < 1.1$). There was an 18-msec word frequency effect for gaze durations; however, it was marginally significant only by participants [$t_1(35) = 1.84, p < .10; t_2(19) = 1.46, p < .20$]. This may have been due to the constraint that there are few prefixed words that are truly high in frequency; thus the difference in frequency between the high and low word frequency conditions was considerably smaller than that for the root frequency set. The picture was similar for the next measure, go-past, since there was only a 16-msec effect [$t_1(35) = 1.31, p < .20; t_2(19) = 1.05, p > .20$]. There was little effect of word frequency on spillover duration, since the low- and high-frequency conditions differed by only 3 msec ($ts < 1$). There was a

slightly larger delayed effect of word frequency than of root frequency, but it was not significant. Second-pass time for high-frequency words was 20 msec less than that for low-frequency words [$t_1(35) = 2.11, p < .05; t_2(19) = 1.30, p > .20$]. The total time measure, which accumulates both early and later effects of word frequency, showed a marginal effect of word frequency; low-frequency words were fixated a total of 34 msec longer than high-frequency words [$t_1(35) = 2.70, p < .025; t_2(19) = 1.66, p < .15$].

We also examined two other indices of word processing: the landing position on the target word and the duration of the fixation prior to the landing on the target word. For both landing position and the duration of the prior fixation, there were virtually no differences between word frequency conditions ($ts < 1$).

Post hoc analyses. Does word length modulate frequency effects? As was indicated earlier, a recent study in Finnish by Bertram and Hyönä (2003) showed larger word frequency effects for short compound words than for long compound words and larger constituent frequency effects for long than for short compound words. On the basis of their finding, post hoc analyses were run to determine whether there was a similar modulation of root frequency and word frequency effects by length for this set of prefixed English words.

Two types of post hoc analyses were run on the Experiment 1 data: t tests and regression analyses. For the t tests, each set of stimuli was divided into two subsets categorized by word length. The target word pairs in the *long word* sets averaged 7.5 or more letters, whereas the target word pairs in the *short word* set averaged 7 or fewer letters. For the *long* root frequency set, gaze durations for the low root frequency words was 75 msec longer than those for the high root frequency words [$t_2(8) = 2.25, p < .06$] but the difference was only 22 msec for the *short* root frequency set [$t_2(6) = 1.48, p < .20$]. The word frequency set showed the opposite pattern. For the *long* word frequency set, word frequency had only a 1-msec effect ($t < 1$), whereas for the *short* word frequency set, the low-frequency words were fixated 42 msec longer than the high-frequency words [$t_2(8) = 2.21, p < .06$]. These findings are thus similar to those in Bertram and Hyönä (2003), since longer prefixed words showed a bigger root morpheme frequency effect than did shorter words, whereas only the shorter prefixed words showed a word frequency effect.

One problem with a dichotomous characterization, such as the one above, is that words of fairly similar lengths are put into opposite groups, weakening the power of the analysis. As a result, both the root frequency and word frequency sets were analyzed with a regression analysis in which word length was entered as the predictor variable for either the root frequency effect of a target word pair or the word frequency effect of a target word pair. The effect of root frequency increased with word length [$t(14) = 4.91, p < .001$], whereas the effect of word frequency decreased with word length [$t(18) = 2.61, p <$

Table 3
Experiment 1: Reading Measures as a Function
of Word Frequency

Reading Measure	Word Frequency	
	High	Low
Probability of skipping target word (%)	3.8	3.0
First-fixation duration (msec)	275	281
Gaze duration (msec)	334	353
Go-past time (msec)	380	396
Spillover duration (msec)	274	277
Second-pass time (msec)	45	64
Total time (msec)	377	411
Initial landing position on target word	2.99	3.01
Time of last fixation prior to fixating target (msec)	239	237

.025]. Thus, the regression analyses clearly indicate that word length modulated both the root frequency and the word frequency effects. We also examined the first-fixation duration data for the word frequency data, since there appeared to be a hint of an effect in the overall data; in fact, length appeared to decrease the size of the effect on first-fixation duration (as with gaze duration), although the effect was not quite significant [$t(18) = 2.00$, $p < .10$]. The slopes of the regression lines were 28 msec per letter for the root frequency effect on gaze duration, -20 msec per letter for the word frequency effect on gaze duration, and -12 msec per letter for the word frequency effect on first-fixation duration.

Summary. Experiment 1 demonstrated that there were large and significant effects of the root frequency of a prefixed word on gaze duration, which is probably the best single measure of word-encoding time in reading. The effect of word frequency on gaze duration, however, was quite a bit smaller, and the overall effect was, at best, marginally significant. However, post hoc tests suggested that both effects were modulated by word length: The root frequency effect was larger for longer words, whereas the word frequency effect was larger for shorter words, which is a pattern analogous to that which Bertram and Hyönä (2003) found for Finnish compound words.

EXPERIMENT 2

Experiment 2 had two purposes. The first was to attempt a replication of Experiment 1—that is, to determine whether one could again obtain a reliable root frequency effect and perhaps obtain a reliable word frequency effect. The second was to examine whether there would be modulation of these effects by word length when there was more careful experimental control of other relevant variables when word length was varied.

Method

Participants. Forty-four native English speakers from the University of Massachusetts community participated in the experiment. Each participant received either course credit or money for his or her participation.

Apparatus. The apparatus was identical to that in Experiment 1.

Materials. Prefixed English words were used as target items. As in Experiment 1, each target word was orthographically transparent and was composed of a prefix plus a free root morpheme. There were four sets of target word pairs, with the words in two of the sets of pairs differing in root frequency and the words in the other two sets differing in word frequency. The two root frequency sets and the two word frequency sets differed in length. The two target words in each pair shared the same prefix and were equal in word length[3] (number of letters) and number of syllables, and across word sets, average frequencies and lengths were matched as closely as possible when appropriate. The mean stimulus characteristics for each set of stimuli are displayed by condition in Table 4, and the stimuli are included in Appendix B. In both the short root frequency and the short word frequency sets, the words were seven or fewer letters, whereas the words in both the long root frequency set and the long word frequency sets were eight or more letters. There were 15 pairs of items in each of the word frequency sets and 14 pairs of items in each of the root frequency sets.

Experiment 2 used the same counterbalancing design as that in Experiment 1. Each pair of target words was embedded in a sentence frame, with a minimum of two words before and after the target word in every sentence. Because the two target words within each pair were placed alternately within the same sentence frame, there were two lists of sentences constructed, so that each participant saw only one word from each pair and each sentence frame only once. In addition, the frequency conditions were counterbalanced across both lists, so that each participant saw items from all eight conditions. This resulted in each participant's seeing 58 experimental sentences. In addition, there were 128 filler sentences, and the order of the 186 sentences seen (58 experimental and 128 filler) was randomized separately for each participant.

To make sure that the differences between the conditions were not caused by a difference in how well the two target words fit into their sentence frames, the experimental sentences were rated by a separate set of participants. Ten participants were asked to read the sentence pairs and rate the relative naturalness of the sentences, as in Experiment 1, where they responded that (1) the sentence containing the low-frequency word was more natural, (2) the sentence containing the high-frequency word was more natural, or (3) the sentences were equally natural. As in Experiment 1, the order of the sentences was balanced, so that for half of the items, the sentence containing the low-frequency target was listed first and, for the other half of the items, the sentence containing the high-frequency target was listed first. Each response was numerically coded as -1, $+1$, or 0, respectively, as in Experiment 1. The mean naturalness difference scores for the four sets were the following: short root frequency, $+0.20$; long root frequency, -0.06; short word frequency, $+0.07$; and long word

Table 4
Experiment 2: Stimulus Characteristics

| Stimulus Characteristic | Root Frequency | | | | Word Frequency | | | |
| | Low | | High | | Low | | High | |
	M	SD	M	SD	M	SD	M	SD
			Short Words					
Mean word frequency	2.50	2.82	1.14	0.95	1.07	1.22	24.27	16.70
Mean root frequency	11.86	6.79	100.10	105.00	76.47	78.80	78.87	69.40
Mean length	6.29	0.47	6.29	0.47	6.33	0.49	6.33	0.49
Mean no. of syllables	2.14	0.36	2.14	0.36	2.27	0.46	2.27	0.46
			Long Words					
Mean word frequency	2.21	2.33	2.43	3.08	2.87	0.80	19.00	10.60
Mean root frequency	13.43	7.76	100.20	90.50	77.07	91.20	71.40	54.50
Mean length	8.79	1.63	8.79	1.63	8.40	0.74	8.40	0.74
Mean no. of syllables	3.14	0.86	3.14	0.86	2.87	0.52	2.87	0.52

frequency, -0.04. A plus value indicates that the sentence with the higher frequency member of the pair was rated as more natural. The matching was almost perfect in three of the four sets; the only set that had any bias was the short root frequency set, in which the sentences with the higher root frequency were judged to be a bit more natural than those with the lower root frequency. However, if anything, this would work against finding a length × root frequency interaction.

Procedure. The procedure was identical to that in Experiment 1, with one exception. In Experiment 2, the participants answered comprehension questions after only 11% of the sentences.

Results

Data loss and exclusion criteria. The data loss and exclusion criteria were identical to those in Experiment 1. In Experiment 2, 3 participants were excluded from the analyses because they exceeded the 5% data loss criterion. Several items were lost due to other errors, but the data loss rate for the final set of 44 data files analyzed remained well below 1%. On average, the participants scored over 97% correct on the comprehension questions, and no participant was excluded on this basis. In addition, after running the experiment, we noticed that there was one word in the long root frequency set (*recreate*) that was a homograph whose intended meaning (to create again) was much lower in frequency than the more usual meaning. As a result, this word and its matched item were excluded from the analyses. (The overall pattern of means was quite similar when this item was included, but the variability was much greater.)

Root frequency. As in Experiment 1, the target words were almost always fixated, and there was little effect of frequency on how often they were skipped (Fs < 1); however, the short words were skipped more often than the long words [$F_1(1,43) = 11.90, p < .005; F_2(1,25) = 8.68, p < .01$; see Table 5]. As in Experiment 1, there was little effect of the frequency of the root morpheme on first-fixation duration (Fs < 1), but there was a suggestion that first-fixation durations for longer words were slightly longer than those for shorter words [$F_1(1,43) = 6.08, p < .025; F_2(1,25) = 1.97, p > .15$].

For gaze duration, the pattern mirrored that in our post hoc analysis in Experiment 1. For the shorter words, the gaze durations for high root frequency words were actually 5 msec longer than those for low root frequency words (ts < 1), but for the longer words, there was a 52-msec advantage for the high root frequency words [$t_1(43) = 3.64,$

$p < .005; t_2(12) = 2.38, p < .05$]. The overall effect of root frequency on gaze duration was not significant in the item analysis [$F_1(1,43) = 8.92, p < .01; F_2(1,25) = 3.26, p < .10$], but the interaction of length with root frequency was significant [$F_1(1,43) = 10.23, p < .005; F_2(1,25) = 4.39, p < .05$]. The overall effect of length was also significant [$F_1(1,43) = 60.09, p < .001; F_2(1,25) = 10.90, p < .005$].[4] The effects in the slightly later measure, *go-past* (which includes all fixations from the time at which the target word is first entered from the left until it is exited to the right), were similar to the gaze effects (see Table 5). Again, there was no reliable root frequency effect for the shorter words (ts < 1), but the 66-msec root frequency effect for the longer words was reliable [$t_1(43) = 3.35, p < .005; t_2(12) = 2.36, p < .05$], as was the length × frequency interaction [$F_1(1,43) = 8.27, p < .01; F_2(1,25) = 5.70, p < .05$]. As with gaze duration, the overall length effect was highly reliable [$F_1(1,43) = 34.97, p < .001; F_2(1,25) = 12.52, p < .005$], but the overall frequency effect was not reliable by items [$F_1(1,43) = 7.22, p < .025; F_2(1,25) = 2.48, p < .15$].[5]

There were only small effects of root frequency or length on measures that assessed processing after the first pass on the target word. For spillover duration, there was virtually no effect of root morpheme frequency for either length (ts < 1) and only the slightest suggestion that it was longer for longer target words [$F_1(1,43) = 1.97, p < .20; F_2(1,25) < 1$]. For second-pass time, there was a main effect of length, with greater second-pass times for longer words [$F_1(1,43) = 11.93, p < .005; F_2(1,25) = 6.47, p < .025$], but no main effect of frequency (Fs < 1). However, there was a suggestion (see Table 5) that there was a *reverse frequency effect* for the shorter words [$t_1(43) = 2.31, p < .05; t_2(13) = 1.76, p < .15$]. The pattern in the total time measure (the sum of gaze duration and second-pass time) was similar to that for gaze duration. There was a main effect of word length [$F_1(1,43) = 55.70, p < .001; F_2(1,25) = 16.03, p < .001$], as well as a root morpheme effect for long words and a frequency × length interaction that were significant only in the participant analysis [$t_1(43) = 2.61, p < .025$, and $t_2(12) = 1.91, p < .10; F_1(1,43) = 9.58, p < .005$, and $F_2(1,25) = 4.20, p < .10$].

We will report analyses of two other measures. The first is the initial landing position on the target word. As can be seen in Table 5, the mean landing position was about half

Table 5
Experiment 2: Reading Measures as a Function of Root Frequency and Word Length

Reading Measure	Short Words		Long Words	
	Low Root Frequency	High Root Frequency	Low Root Frequency	High Root Frequency
Probability of skipping word (%)	4.7	5.0	0.8	1.2
First-fixation duration (msec)	266	265	273	278
Gaze duration (msec)	321	326	413	361
Go-past time (msec)	346	359	458	394
Spillover duration (msec)	251	252	261	259
Second-pass time (msec)	25	41	67	63
Total time (msec)	344	361	481	425
Initial landing position on target word	2.66	2.55	3.06	3.19
Time of last fixation prior to fixating target (msec)	233	251	238	235

a character further into the word when it was long than when it was short [$F_1(1,43) = 20.8, p < .001; F_2(1,25) = 7.71, p < .01$]. This suggests that the root of the word may have been in a slightly better location for processing for the long words. However, there was no effect of root frequency and no interaction between frequency and length ($Fs < 1$). The second measure is the duration of the fixation prior to the reader's landing on the target word on the first pass. There was a suggestion of an effect that was a bit strange and hard to interpret (see Table 5). This fixation was actually 7 msec shorter for the low root frequency words [$F_1(1,43) = 5.27, p < .05; F_2(1,25) = 1.72, p = .20$], and there was a suggestion of a frequency × length interaction [$F_1(1,43) = 6.87, p < .05; F_2(1,25) = 4.10, p < .10$], indicating that most of the difference was in the short word condition. When the short words were analyzed separately, the 19-msec difference was almost significant [$t_1(43) = 3.51, p < .005; t_2(13) = 2.08, p < .10$]. (The absolute difference between the short and the long word conditions is not interpretable, since the prior text was different in the two conditions.) This suggests that perhaps some processing of the root occurred before the target word was fixated in the short word condition. If this effect is real, it seems to be a puzzle as to why this fixation duration would be shorter for words with low root frequencies, however.

Word frequency. Target words in all four conditions were fixated on a high percentage of the trials, but there was a clear effect of word length and a marginal effect of word frequency (see Table 6). The short words were skipped 5.4% more often than the long words [$F_1(1,43) = 15.42, p < .001; F_2(1,28) = 17.01, p < .001$], and the high-frequency words were skipped 1.7% more often than the low-frequency words [$F_1(1,43) = 3.11, p < .10; F_2(1,28) = 3.40, p < .10$]. (The interaction was negligible; $Fs < 1$.) Unlike root frequency, word frequency had a substantial effect on first-fixation durations, particularly for the short word set. The 15-msec main effect of word frequency was significant [$F_1(1,43) = 14.63, p < .001; F_2(1,28) = 8.77, p < .01$], as was the 19-msec frequency effect for the short words [$t_1(43) = 3.42, p < .005; t_2(14) = 2.90, p < .025$], but the 10-msec frequency effect for the long words was not significant by items [$t_1(43) = 2.35, p < .025; t_2(14) = 1.34, p > .20$]. Neither the length

main effect nor the length × frequency interaction was close to significant (all $Fs < 2$).

There was a 33-msec main effect of word frequency on gaze duration [$F_1(1,43) = 15.30, p < .001; F_2(1,28) = 13.06, p < .005$]. The 38-msec word frequency effect for short words was also significant [$t_1(43) = 3.30, p < .005; t_2(14) = 3.42, p < .005$], but the 28-msec word frequency effect for the long words was only marginal [$t_1(43) = 3.08, p < .005; t_2(14) = 1.81, p < .10$]. The frequency × length interaction was not close to significant ($Fs < 1$).[6] The 43-msec main effect of length was also significant [$F_1(1,43) = 48.45, p < .001; F_2(1,28) = 12.04, p < .005$]. In the go-past measure, the interaction of frequency and length seemed somewhat stronger, since the word frequency effect for the short words was 50 msec [$t_1(43) = 4.01, p < .001; t_2(14) = 3.95, p < .005$], whereas the 14-msec word frequency effect for the long words was not close to significant ($ts < 1$). However, the interaction was not significant by items [$F_1(1,43) = 4.31, p < .05; F_2(1,28) = 2.45, p < .15$]. As with gaze duration, the main effects of both word frequency and length were significant [$F_1(1,43) = 9.00, p < .005$, and $F_2(1,28) = 7.02, p < .025; F_1(1,43) = 33.38, p < .001$, and $F_2(1,28) = 12.59, p < .005$, respectively].

There were also significant effects of both word frequency and length on later measures of processing. For spillover duration, there was an 11-msec word frequency main effect [$F_1(1,43) = 6.32, p < .025; F_2(1,28) = 11.42, p < .005$], but virtually no effect of length and no interaction between the two ($Fs < 1$). For second-pass time, the 42-msec word frequency effect for short words was significant [$t_1(43) = 4.25, p < .001; t_2(14) = 2.61, p < .025$], whereas the 12-msec effect for long words was not close to significant ($ts < 1.1$). However, the frequency × length interaction was not significant [$F_1(1,43) = 3.61, p < .10; F_2(1,28) = 1.30, p > .20$]. The total time measure again indicated that the word frequency effect was greater for short words—75 msec for the short words [$t_1(43) = 5.37, p < .001; t_2(14) = 4.03, p < .005$] versus 44 msec for the long words [$t_1(43) = 3.58, p < .005; t_2(14) = 1.46, p < .20$]—but the interaction again was not significant [$F_1(1,43) = 2.76, p < .15; F_2(1,28) = 1.10, p > .20$].

As with the root frequency set, we also examined the initial landing position on the target word and the duration

Table 6
Experiment 2: Reading Measures as a Function of Word Frequency and Word Length

Reading Measure	Short Words		Long Words	
	Low Word Frequency	High Word Frequency	Low Word Frequency	High Word Frequency
Probability of skipping word (%)	6.6	8.2	1.0	3.0
First-fixation duration (msec)	269	250	270	260
Gaze duration (msec)	322	284	359	331
Go-past time (msec)	351	301	387	374
Spillover duration (msec)	267	254	268	259
Second-pass time (msec)	67	26	69	57
Total time (msec)	382	307	429	385
Initial landing position on target word	2.55	2.54	2.93	2.78
Time of last fixation prior to fixating target (msec)	215	227	226	232

of the fixation prior to the reader's landing on the target word on the first pass (see Table 6). As with the root frequency items, the initial fixation location on the target word was a bit further into the target word [about 0.3 characters; $F_1(1,43) = 8.58$, $p < .01$; $F_2(1,28) = 5.33$, $p < .05$], and there was no effect of word frequency and no interaction between frequency and length ($Fs < 1$). For the prior fixation, there was again a suggestion of a small effect in the "wrong" direction. That is, this fixation duration was 9 msec shorter for the low-frequency words [$F_1(1,43) = 4.51$, $p < .05$; $F_2(1,28) = 2.16$, $p < .20$]. The effect appeared to be somewhat bigger for the short words, but the interaction was not close to significant [$F_1(1,43) = 1.27$, $p > .20$; $F_2 < 1$]. (Again, note that absolute differences between short and long words were not easily interpretable, because the text prior to the target word was different in the two conditions.) As with the root frequency set, we are mystified by these results if they are real and not due to a subtle artifact that we have failed to unearth.

Summary. The general pattern of data matched the predictions generated from Experiment 1. In Experiment 2, there was a clear effect of root frequency on the longer words for both gaze duration and go-past time and a significant interaction for both these measures of length and frequency. Moreover, there was clearly no effect of the root morpheme frequency for the shorter words. For word frequency, the opposite pattern—a larger word frequency effect for the shorter words—was obtained for gaze duration, go-past, and total time measures, but the interaction of frequency with word length was not significant in the item analysis for any of these measures. However, unlike in Experiment 1, there was a reliable effect of word frequency on all of these measures, which was a reflection of the fact that although the word frequency effect was not significant for the long words, there still appeared to be a reduced effect.

Experiment 2 also indicated some interesting differences in the pattern of effects for the root frequency and word frequency manipulations. First, there were reliable word frequency effects on first-fixation duration but no root frequency effects. Second, there were reliable word frequency effects on spillover duration and second-pass time but no root frequency effects. These differences mirrored the pattern observed in Experiment 1 (with the exception of the spillover duration).

FURTHER ROOT FREQUENCY AND WORD FREQUENCY ANALYSES

All of the analyses and conclusions above about the roles of root and word frequency were dependent on particular measures of root and word frequency, and since any particular measures are likely to be fallible, we undertook subsidiary analyses in order to determine whether our conclusions would be substantially different if other measures were used. Of course, none of these measures can pretend to be anything other than a guide for how frequently any given participant has seen a particular word or a particular word pattern in his or her lifetime.

There are two separate issues. The first concerns the norms used for determining word frequency (and hence, any calculation of root frequency), and the second concerns the best measure for computing root frequency from these norms. We will deal with the two issues in turn. The first is particularly relevant to our conclusions about the effects of root frequency on processing time, because the logic of our experiments is that root frequency has been varied while the surface frequency of the prefixed words has been equated. If, however, the surface frequency of the prefixed words has not been held constant but, instead, the words with higher root frequency have higher surface frequency, one can not unambiguously claim to have found an effect of root frequency. Hence, we decided to determine whether our sets of stimuli in which root frequency was manipulated were, in fact, equated on norms other than those of Francis and Kučera (1982). (We will simply state for now that, by all other measures, these sets of stimuli did indeed differ substantially on root frequency.)

There are a variety of other measures that are in common use. The two we chose to investigate, the *U* measure (Zeno, Ivens, Millard, & Duvvuri, 1995) and the written word frequency measure (per million) from the CELEX corpus (Baayen, Piepenbrock, & Gulikers, 1995), seemed to be representative, and analyses done with other measures from these corpora were virtually identical to the ones we will present below. In fact, according to these two corpora, our *root frequency* sets were not equated for mean word frequency (see Table 7). As with the Francis–Kučera (1982) values, there was a bias for the *high root* set to have a somewhat higher mean word frequency than did the *low root* set in Experiment 1 (see Table 1) and for the bias to be in the opposite direction for the short root frequency set in Experiment 2 (see Table 3). The principal difference was that for the long root frequency set, for these corpora, the high root set now also had a higher mean word frequency, whereas, according to the Francis and Kučera norms, the two sets were quite well equated (see Table 3).

We conducted a series of regression analyses to determine whether these differences in word frequency (according to these other two corpus measures) could have been responsible for the differences in root frequency that we reported. To simplify exposition, we will confine ourselves to analyses of gaze duration. Our general method was to take the difference in mean gaze durations for the items in each pair of items and determine whether the difference in word frequency for that pair of items, as determined by either Zeno *U* or CELEX (see Table 7), could have predicted the observed difference in gaze duration. That is, in the regression analysis, we were interested primarily in the intercept of the regression (i.e., the best prediction for the difference in mean gaze duration when the difference in word frequency was zero). However, in Experiment 1, there were two other variables that were likely to be of importance in the regression analysis. First, as was mentioned earlier, not all the pairs of items in Experiment 1 were of equal length (although the mean length of the pairs was equated). As a result, we thought it best to remove this source of variance from the regression. More

Table 7
Alternate Measures of Mean Word Frequency for the Root Frequency Effects

	Experiment 1			Experiment 2					
				Short Words			Long Words		
Measure	Low Root Frequency	High Root Frequency	Difference	Low Root Frequency	High Root Frequency	Difference	Low Root Frequency	High Root Frequency	Difference
Zeno et al. (1995) *U*	0.69	2.31	1.62	1.02	0.49	−0.53	0.54	2.61	2.07
CELEX written word frequency	2.18	3.69	1.51	2.27	1.06	−1.21	2.10	3.68	1.58

important, the words varied considerably in length, and this was found to have a substantial effect on the root frequency effect in our prior analyses. As a result, the mean length of each of the pair of words was entered into the regression equation, except that the variable actually used in the regression equation was the mean length of the pair of words minus the mean length of all the stimuli (which was 7.9 letters), so that the intercept of the regression equation would be interpretable: It would be the best prediction of the true root frequency effect for prefixed words of this mean length.

The regression analyses using the *U* and CELEX measures came out quite similarly. The intercepts of the analyses, which are interpretable as the estimated root frequency effect on gaze duration for a word of 7.9 characters with the word frequency measure equated, were 47 and 41 msec for the *U* and CELEX measures, respectively [$t(12) = 3.31, p < .01$, and $t(12) = 3.05, p < .01$, respectively]. (The mean gaze duration over items was 52 msec.) As in our regression analyses reported in Experiment 1, this root frequency effect was substantially affected by length, with slopes of 30 msec per character in both analyses [$t(12) = 4.79, p < .001$, and $t(12) = 5.08, p < .001$]. In both analyses, the difference in word frequency measure appeared to have some effect in the predicted direction (slopes of 3.4 and 7.5 msec per unit difference in frequency measure). Only the latter effect was close to significant [$t < 1; t(12) = 1.82, p < .10$]. The slopes for differential length were also not significant [$t(12) = 1.12, 1.05, ps > .20$].

In Experiment 2, we merely attempted a regression of the difference in gaze duration with the difference in the word frequency measure as the predictor, since the two sets were relatively homogeneous with respect to length and all the word pairs had equal numbers of letters. Unlike in Experiment 1, these analyses affected the root frequency effects a bit. For the short words, the estimated sizes of the root frequency effect on gaze duration were 12 and 11 msec, using differences in *U* and CELEX, respectively ($ts < 1$), and for the long words, the estimated sizes of the root frequency effect on gaze duration were 34 and 45 msec [$t(12) = 1.26, p > .20$, and $t(12) = 1.75, p < .20$]. (Again, the word pair in the long word set that included the homograph *recreate* was excluded from the analyses.) These values are somewhat different from the item means for the differences in gaze durations for the short and long word sets (2 and 60 msec, respectively) and appear to reflect the fact that the effect of the uncontrolled word frequency (using these

measures) had some effect. The slopes of the regression on the difference in word frequency (in milliseconds per unit frequency measure) were bigger than those in Experiment 1; however, in none of the analyses was the slope of the regression coefficient significant. For the short words, the slopes were 18.6 and 7.9 [$t(12) = 1.44, p > .10$, and $t(12) = 1.57, p > .10$], and for the long words, they were 12.3 and 9.3 [$t(12) = 1.87, p < .10$, and $t(12) = 1.40, p > .20$].

In sum, the pattern of data in these analyses seems quite similar to the pattern reported in the main analyses in Experiments 1 and 2. The chief difference is that the effect of length on the difference in gaze duration was attenuated a bit in Experiment 2.[7] However, all the analyses indicated that there was a root frequency effect on gaze duration for longer prefixed words. Admittedly, the intercepts in the regression analyses in Experiment 2 were not significant, but these analyses were not very powerful, not having many degrees of freedom. In addition, in these experiments, there were huge numbers of constraints on the equating of words on various measures and simultaneously ensuring that a pair of words one had selected fit naturally into the same sentence. Thus, one often buys an increase in apparent power by making the materials less natural.

We now will turn to the question of alternative assessments of the equating of root frequency. In the original analyses, we simply used the frequency of the root as an isolated word. It was of prime importance to be sure that the word frequency effects we observed would hold up with different measures of root frequency in the equating of the sets. In the first analysis, we used the Francis and Kučera (1982) norms and used the sum of the frequencies of all the words that contained the root (except that we did not include inflections). As can be seen in Table 8, the word frequency sets were reasonably well equated on this new variable in Experiment 2, but not in Experiment 1.

For Experiment 1, we did a regression analysis parallel to that on the root frequency words, where we took the difference in gaze durations for a word pair and used the difference in the summed root frequency, the mean length of the pair minus the mean length of all the pairs (7.6 letters), and the difference in length of the word pair as predictors. In this analysis, we found that even though the summed root frequency was not well matched, this variable had virtually no predictive power, since the slope was 0.016 msec per unit difference in frequency per million ($t < 1$). The intercept in the regression equation was 14 msec [$t(16) = 1.06, p > .20$], which was only a bit smaller than the item

Table 8
Alternate Measures of Mean Root Frequency for Word Frequency Effects

	Experiment 1			Experiment 2					
				Short Words			Long Words		
Measure	Low Word Frequency	High Word Frequency	Difference	Low Word Frequency	High Word Frequency	Difference	Low Word Frequency	High Word Frequency	Difference
Summed Francis & Kučera (1982)									
root frequency	391	663	272	168	280	112	172	210	38
Summed CELEX root frequency	354	678	324	185	294	109	152	201	49

mean difference between the low word frequency and high word frequency conditions (19 msec). However, as in the main analysis in Experiment 1, this difference was strongly modulated by word length, so that the slope with respect to word length was −20 msec per letter [$t(16) = 2.36, p < .05$]. In Experiment 2, we did a simpler analysis, in which the only predictor was the difference in the summed root frequency. For the short words, the intercept was 44 msec [$t(13) = 2.73, p < .01$], and for the long words, the intercept was 28 msec [$t(13) = 1.79, p < .10$], which were only slightly different from the uncorrected item means (50 and 30 msec, respectively). In the two analyses, the slopes were 0.003 and 0.045 msec per unit difference in root frequency (ts < 1). In the case of the short words, there may be this small a coefficient because there appears to be little effect of root frequency for shorter words from the root frequency analyses. For the longer words, the slope was a bit bigger, but there was a smaller difference in the summed root frequency for the two sets (two items produced nearly all the difference in summed root frequency).[8]

Given that the word frequency analyses of the root frequency manipulation came out quite similarly with the CELEX and Zeno norms, we redid the analyses in the above paragraph with only the summed root frequencies computed from the CELEX corpus. The results of the analyses were quite similar to the ones in the preceding paragraph. In Experiment 1, the intercept in the regression equation was 13 msec ($t ≈ 1$), the slope for length was −19 msec per letter [$t(16) = 2.35, p < .05$], and the slope for the difference in the CELEX index was 0.017 msec per unit difference in frequency per million ($t < 1$). In Experiment 2, for the short words, the intercept was 46 msec [$t(13) = 2.81, p < .05$], and the slope for the difference in the CELEX index was 0.050 msec per unit difference ($t < 1$); for the long words, the intercept was 27 msec [$t(13) = 1.57, p > .20$], and the slope for the difference in the CELEX index was 0.050 msec per unit difference ($t < 1$).

In sum, despite the significant imbalance of summed root frequency between the high and the low word frequency sets (especially for Experiment 1), the effects in the regression analyses using both summed root frequency values were quite similar to those in the primary analyses reported earlier, both in absolute size and in the reliability of the results. As in the primary analysis in Experiment 1, the overall effect of word frequency was only about 15 msec and was not significant, but there was a highly significant effect of length, indicating that there

was a large word frequency effect for the shorter words. Similarly, the regression analyses in Experiment 2 using both summed root frequency values indicated that there was a large and significant word frequency effect for the shorter words and a smaller and nonsignificant word frequency effect for the longer words. These analyses suggest that summed root frequency has little predictive power beyond the frequency of the root in isolation for gaze durations.

GENERAL DISCUSSION

The present experiments clearly demonstrated that the root morphemes of prefixed words in English are involved in the encoding process, at least for longer prefixed words. In Experiment 1, a significant 54-msec effect of root morpheme frequency was observed for the gaze duration on the critical prefixed word (47 and 41 msec in the two regression analyses using the Zeno and CELEX databases, respectively); however, post hoc tests indicated that the size of the effect was modulated by word length. Experiment 2 replicated this effect, but only for longer words. For the long word set, there was a significant 52-msec root frequency effect, whereas there was actually a −5-msec root frequency effect for the shorter words. In the analogous regression analyses, these effects were less divergent—34 and 45 msec for the long words and 12 and 11 msec for the short words—and the effects for the long words were not significant. We used the regression equation derived in our later regression analyses (i.e., with the values corrected for differences in word frequency) to estimate what the predicted size of the root frequency effect would have been in Experiment 1 for words of a length equal to that of our short words in Experiment 2 (6.3 letters) and for words equal to our long words in Experiment 2 (8.7 letters). The predicted values were −18 and 54 msec, which are close to the uncorrected values in Experiment 2 and not that divergent from the corrected ones. The finding of a root frequency effect in gaze duration, but no root frequency effect in first-fixation duration for longer words, is thus quite consistent with the findings of Beauvillain (1996), who found a root frequency effect on gaze duration (but not on first-fixation duration) for long French words (averaging about 10 characters in both experiments) in a task involving two isolated words.

The data also indicated that word frequency affected encoding time, and the size of this effect was also affected by the length of the word. In Experiment 1, there was only

a 19-msec effect of word frequency on gaze duration, which was not significant; however, a post hoc analysis indicated that there was a significant modulation of this effect by word length. In Experiment 2, in contrast, the 36-msec effect of word frequency was highly reliable, but the 16-msec difference in the size of this effect between the long word set and the short word set (44 vs. 28 msec) was not significant. (None of these analyses was significantly affected by using the summed root frequencies as the measure of root frequency rather than the frequency of the root in isolation.) When we tried to predict the size of the Experiment 2 effects from the Experiment 1 regression equation (analogous to the prediction in the paragraph above), the prediction for the short words was 44 msec, and that for the long words was 4 msec. Thus, the results seemed quite consistent across both experiments for the longer words but a bit different for the shorter words.

It is also worth mentioning that there were significant effects of word frequency on first-fixation duration in Experiment 2 (15 msec overall and 19 msec for the shorter words), and although the 6-msec word frequency effect on first-fixation duration in Experiment 1 was not significant, the regression analyses indicated that it would have been about 15–20 msec for words equal in length to the shorter words in Experiment 2. In contrast, there were no reliable root frequency effects on first-fixation duration (6 msec in Experiment 1, 1 msec for the short words in Experiment 2, and −5 msec for the long words in Experiment 2). Thus, it appears that, if anything, word frequency effects occur earlier than root frequency effects, especially for shorter prefixed words.

The pattern we observed in gaze duration, as was mentioned earlier, was analogous to the pattern observed by Bertram and Hyönä (2003) for Finnish compound words. The pattern of their two experiments also differed somewhat, since the root frequency effect in Experiment 1 for their shorter words was 20 msec (but smaller than that for their longer words), whereas in Experiment 2, there was no root frequency effect. Similarly, their Experiment 2 showed a smaller (20-msec) word frequency effect for the longer words than for the shorter words, whereas Experiment 1 showed virtually no word frequency effect for the longer words. Thus, both their experiments and ours show definite compositional effects for longer words and small and unreliable compositional effects for shorter words; in addition, both sets of experiments show definite whole-word effects for shorter words and small and unreliable whole-word effects for longer words.

What can one infer from this pattern of results about how these prefixed words are processed? We think that the most parsimonious explanation for the data is that there is a dual-route mechanism in which a compositional access mechanism and a whole-word *direct* mechanism operate in parallel. (We will leave open the issue of whether these processes interact.) First, let's consider the shorter words. We found a word frequency effect and little or no evidence of root frequency effects for these shorter words. Given a dual-route model, the most parsimonious mechanism

for accounting for this data pattern is that both routes are activated but that the whole-word route is fast enough that there is little or no effect of the compositional route on access of the word and, hence, on the eye movement pattern. We think that this would be a reasonable inference even if there were absolutely no root frequency effect. The alternative—that the compositional route is "silent" for these shorter words—seems less parsimonious. That is, it would seem to require an executive mechanism that would decide that the compositional process should be turned off when a word is shorter than x letters. Such a mechanism would seem to be unwieldy and suboptimal; among other things, how does it decide what x is?

For the longer words, since there is clear evidence of root morpheme frequency effects, the compositional route appears to be involved. The fact that the two experiments differ to some extent as to whether there are word frequency effects for the longer words makes it somewhat more difficult to make strong conclusions about how these words are processed. However, as was argued above, a model in which one or the other process is turned off, given a preliminary computation of length, seems implausible. Instead, it would seem more natural that both processes run in parallel but that the compositional process largely dominates for the longer words. Presumably, the primary reason for this is that it gets harder to process the word in parallel in a single fixation when it is longer and, hence, the whole word process slows down more than the componential processes. A second factor may be that, as our landing position analyses indicated, the initial fixations in the longer words are further into the word, which may be a relatively better position from which to analyze the root and a relatively worse position from which to analyze the entire word.

One theoretical framework that can account for both overall root frequency and word frequency effects is a variant of the two-stage decomposition model of Taft and Forster (1976). In this model, access to a complex word has two stages (actually, three in the case of prefixed words, since there is a hypothesized preliminary *prefix-stripping* stage that occurs before the two main stages). After the prefix has been stripped, the root morpheme is looked up in Stage 1. The duration of this stage is presumably a function of the frequency of the root morpheme, and hence, this is the posited source of the root frequency effects. The next stage consists of looking up the whole word in the file drawer of words that contain this root morpheme. The duration of this stage is sensitive to the relative frequency of the actual word presented to the other words in the file drawer. If one holds root frequency constant, this relative frequency should be higher, the higher the frequency of the word. Thus, according to this model, word frequency effects should accrue in the second stage of processing. Although the model does account for both types of frequency effects, we do not think that it accounts for our data very well. First, it is not clear how such a model would account for the modulation of the effects due to word length. Second, it does not appear

that the time course of effects is accounted for very well either. That is, there appear to be early effects of word frequency (in skipping probability and in first-fixation duration), whereas there are no such early effects of root frequency. Furthermore, the pattern of data appears to be at variance not only with this particular model, but also with the predictions of a whole class of similar two-stage models in which the second stage is something other than the serial look-up stage of Taft and Forster but, instead, a more parallel competition process (e.g., Caramazza et al., 1988; Lukatela et al., 1980). That is, these other models also posit two serial stages, with word frequency entering in only at the second stage. It might be possible to save these models by positing that the preliminary prefix-stripping stage is influenced by word frequency, but then that would seem to be a very different model. Furthermore, Kambe (2004) found no morphological preview benefit for English words, a finding that does not support early prefix stripping.

Taft's (2004) more recent three-level model has the same problems accounting for the data as the two-step models. Although Taft's (2004) activation framework predicts a combination of root frequency and whole-word frequency effects, it cannot account for the time course of our effects. Specifically, it cannot account for the early word frequency effects, because it predicts only later effects of word frequency during the final step when the two components are integrated. Second, as with the two-stage models, Taft's (2004) three-step activation framework does not explain the modulating effect of length on the effects of word and root frequency.

The alternative conception for a system that produces both word frequency and root frequency effects is, of course, some sort of dual-process model, the simplest version of which is a parallel *race* model (e.g., Bertram, Schreuder, & Baayen, 2000; Pollatsek et al., 2000; Schreuder & Baayen, 1995). In a race model, there are two routes to identification of a word: a direct look-up route and a componential route. Which route predominates depends on the relative speed of the two processes. In general outline, such a model seems fairly consistent with the data. First, let's consider the effect of word length. As words get longer, it presumably gets harder to process the entire string of letters as a whole, since some of the letters will not be sufficiently visible on any given fixation. Moreover, it might be hard to align the partial computations of component letters from successive fixations. In contrast, if the prefix can be quickly processed, one may then be able to subsequently process the root morpheme more easily as a smaller entity. Thus, it seems natural that some version of such a model could predict that even though both whole-word processing and componential processing will be slowed down as word length increases, whole-word processing will be hurt more by increases in word length. On the other hand, if processing the prefix and finding the beginning of the root morpheme takes time, componential processing is plausibly slower than

whole-word processing for shorter words. Moreover, if the prefix needs to be processed first in the componential process, it also makes sense that whole-word effects would appear earlier than root frequency effects.

Obviously, such a model would need a lot of "filling in" to be a really satisfying explanation, but such a sketch of a model seems like a reasonable approximation. One reservation we have about a simple parallel race model is that it may underpredict the size of root frequency and word frequency effects. That is, in such a model, the two routes are in a race, and the fastest route wins the race (and thereby makes the time that the slower route takes largely irrelevant to word-encoding time).[9] One attempt to quantitatively fit such a race model to explain component morpheme frequency effects and whole-word frequency effects for Finnish compound words was, to some extent, a failure, since the assumptions about the time to encode morphemes or words to get a reasonable fit of the data seemed implausible (Pollatsek, Reichle, & Rayner, 2003). The pattern of results observed there, admittedly, was somewhat different from the data here, since both very large component frequency effects and very large word frequency effects were observed for long Finnish compound words. Here, in contrast, one or the other effect appeared to predominate. However, in Experiment 2, since both the component frequency effect and the word frequency effects were sizable for the long words (about 40 and 30 msec, respectively), it might also be difficult to model this with a parallel race model. Thus, we suspect that a more complex model, in which the componential and direct look-up processes interact, is more likely to be what is occurring.

There is also a plausibility argument against the processes being independent. That is, how is the meaning of the complex word obtained when the componential process "wins" when the meaning is not deducible from the meanings of the components? Presumably, given such a race model, there should be some sort of *double-take*, where the direct look-up process is consulted to find the appropriate meaning. In the case of long Finnish compound words, however, there was no effect of transparency on the gaze duration on these words, even though there was a clear effect of the frequency of the first morphemic constituent of the compound (Pollatsek & Hyönä, 2005). (This was the case despite the fact that compounding in Finnish is quite productive, so that novel compounds, for which meaning has to be extracted from a componential process, are typically encountered many times a day by Finnish readers.)

We did a post hoc analysis on the data in Experiment 2 to see whether transparency differences modulated the word and root frequency effects we observed, and the analysis revealed no consistent transparency effects. First, if transparency was playing a significant role in constructing the meaning for a word, presumably the difference between high-root and low-root pairs should be reduced when the high-root member of a pair is opaque and the low-root

member is transparent. For the long root frequency set, both words in all but one of the pairs were semantically transparent, so there were no real data. For the short root frequency set, the pattern of differences was strange: The root frequency effect was negative (-23 and -29 msec, respectively), both for the four pairs in which the high-frequency member was opaque and the low-frequency member was transparent and for the two pairs in which the transparency difference was the opposite. For the word frequency pairs, the prediction is less clear, although one might expect that when the high-frequency member of the pair is opaque, it would reduce the frequency effect. This pattern occurred for the short word frequency set (24 msec vs. the overall 51-msec item mean), but the opposite pattern occurred for the long word frequency set (44 msec vs. an overall item mean of 30 msec). This analysis thus provides no clear evidence that transparency is playing an important role in the encoding of prefixed words. As a result, we think a dual-route model is the best explanation of our data, but one in which the component process "co-operates" with the direct look-up process and the meaning is virtually always accessed from the lexical entry for the word (if there is one).

REFERENCES

ALEGRE, M., & GORDON, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory & Language*, **40**, 41-61.

ANDREWS, S., MILLER, B., & RAYNER, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in the mousetrap. *European Journal of Cognitive Psychology*, **16**, 285-311.

BAAYEN, R. H., PIEPENBROCK, R., & GULIKERS, L. (1995). *The CELEX lexical database* [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

BEAUVILLAIN, C. (1996). The integration of morphological and whole-word form information during eye fixation on prefixed and suffixed words. *Journal of Memory & Language*, **35**, 801-820.

BERTRAM, R., & HYÖNÄ, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory & Language*, **48**, 615-634.

BERTRAM, R., SCHREUDER, R., & BAAYEN, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy and productivity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 489-511.

BUTTERWORTH, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production* (Vol. 2, pp. 257-294). London: Academic Press.

CARAMAZZA, A., LAUDANNA, A., & ROMANI, C. (1988). Lexical access and inflectional morphology. *Cognition*, **28**, 297-332.

DEUTSCH, A. F., FROST, R., & FORSTER, K. I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1238-1255.

DOMINGUEZ, A., CUETOS, F., & SEGUI, J. (1999). The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research*, **28**, 485-498.

FRANCIS, W. N., & KUČERA, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

GIRAUDO, H., & GRAINGER, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language & Cognitive Processes*, **15**, 421-444.

HYÖNÄ, J., & POLLATSEK, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1612-1627.

JUHASZ, B. J., STARR, M. S., INHOFF, A. W., & PLACKE, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology*, **94**, 223-244.

KAMBE, G. (2004). Parafoveal processing of prefixed words during eye fixations in reading: Evidence against morphological influences on parafoveal preprocessing. *Perception & Psychophysics*, **66**, 279-292.

LUKATELA, G., GLIGORIJEVIĆ, B., KOSTIĆ, A., & TURVEY, M. T. (1980). Representation of inflected nouns in the internal lexicon. *Memory & Cognition*, **8**, 415-423.

NISWANDER, E., POLLATSEK, A., & RAYNER, K. (2000). The processing of derived and inflected words during reading. *Language & Cognitive Processes*, **15**, 389-420.

POLLATSEK, A., & HYÖNÄ, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language & Cognitive Processes*, **20**, 261-290.

POLLATSEK, A., HYÖNÄ, J., & BERTRAM, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 820-833.

POLLATSEK, A., REICHLE, E., & RAYNER, K. (2003). Modeling eye movements in reading: Extensions of the E-Z reader model. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 361-390). Amsterdam: Elsevier.

RAYNER, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372-422.

SCHREUDER, R., & BAAYEN, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131-154). Hillsdale, NJ: Erlbaum.

SEIDENBERG, M. S., & MCCLELLAND, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, **96**, 523-568.

TAFT, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, **7**, 263-272.

TAFT, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology*, **57A**, 745-765.

TAFT, M., & FORSTER, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning & Verbal Behavior*, **14**, 638-647.

TAFT, M., & FORSTER, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning & Verbal Behavior*, **15**, 607-620.

ZENO, S. M., IVENS, S. H., MILLARD, R. T., & DUVVURI, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science Associates.

NOTES

1. The root frequencies of the high and low word frequency conditions were equated by category (high/low), and the average root frequencies for both categories were quite high; however, the absolute numbers did differ. Therefore, the word frequency effect was retested in Experiment 2, for which main effects of word frequency were found, confirming an effect of word frequency for prefixed English words.

2. Unlike gaze duration, which counts only trials on which the target word is fixated on the first pass, mean second-pass duration counts all trials. Thus, on trials on which there are no regressions to the target word, second-pass time is scored as zero.

3. There was one exception. One of the 15 word frequency pairs differed in length by one character. One member was 13 letters, and the other was 14 characters.

4. When the homographic item *recreate* was included in the gaze duration analyses, there was a 42-msec frequency effect for the long word set [$t_1(43) = 3.28, p < .005; t_2(13) = 1.82, p < .10$]. Both the overall effect of root frequency [$F_1(1,43) = 6.40, p < .025; F_2(1,26) = 1.83, p < .20$] and the interaction of length with root frequency [$F_1(1,43) = 8.13, p < .01; F_2(1,26) = 2.66, p < .15$], although significant in the participant

analysis, failed to reach significance in the item analysis (largely because of the homographic item).

5. When the homographic item *recreate* was included for go-past, the results were quite similar. There was a 59-msec frequency effect for the long word set [$t_1(43) = 3.48, p < .005; t_2(13) = 2.37, p < .05$], as well as a reliable interaction of length with root frequency [$F_1(1,43) = 8.17, p < .001; F_2(1,26) = 5.46, p < .05$], but the overall effect of root frequency failed to reach significance in the item analysis [$F_1(1,43) = 7.10, p < .05; F_2(1,26) = 2.24, p < .15$].

6. As in the root frequency set, there was one target pair that contained a homograph (*invalid*). The item analysis for gaze duration, go-past, and total time were rerun without the homographic item, but the differences between the analyses with and without the homograph were negligible.

7. The length effect in these regression analyses on the go-past measure was somewhat larger. The root frequency effects for the short words were −4 and −5 msec (ts < 1) for the Zeno U and CELEX analyses, respectively, and for the long words were 37 and 42 msec [$t(12) = 1.30, p > .20$, and $t(12) = 1.66, p < .20$].

8. We should also comment that, unsurprisingly, the mean root frequency values (per million), using summed root frequencies, were substantially different for the high and low root frequency pairs. In Experiment 1, the summed root frequencies were 22 and 303. In Experiment 2, using the Francis and Kučera (1982) norms, for the short words, the mean summed root frequencies were 56 and 226 for the low and high root frequency stimuli, and for the long words, the mean summed root frequencies were 60 and 319 for the low and high root frequency stimuli.

9. The time that the loser takes is not completely irrelevant if one assumes that there are random variations in the times for each process (which is the usual assumption in such models); however, the effect of the "losers" not always losing is usually fairly small.

APPENDIX A
Experiment 1 Stimuli

Root Frequency Set [Low/High Frequency]
Everyone's life was in [disarray/disorder] after the hurricane.
The tired detective was about to [disrobe/disarm] when he heard a knock at the door.
Zeus was an [immortal/immoral] god who often had affairs with his worshipers.
Keith couldn't identify the [inanimate/indistinct] object in the distance.
Last meeting's [inarticulate/inappropriate] speaker will certainly not be asked to return.
Jamie's friends noticed she had become [insecure/insincere] since joining a sorority.
Martha witnessed the robbery when she stopped to [refuel/refill] her gas tank.
The fruit was [unripe/unhurt] because the cellar was cold.
Lucy told the jury how her boss had tried to [defraud/degrade] her at work.
Karen found her apartment in [disarray/disorder] after the burglary.
Ben failed to identify the [inaudible/infrequent] tone in the background.
The latest movie was about an [immortal/immoral] vampire who killed movie directors.
Lyn asked her kids to [unfold/unpack] the linens in preparation for the holiday dinner.
The politician's [indiscriminate/inappropriate] behavior was the major topic of discussion.
Mike needed to [unhook/unlock] the rusty latch before he could open the old trunk.
Richard attempted to [refuel/renew] himself by swimming three times a week.

Word Frequency Set [Low/High Frequency]
Although she tried to hide it, Jill's [distrust/dislike] for Ed was obvious to everyone.
Jason's only hope was to [disable/discover] the bomb in time.
The gossip columnist's description of the [imperfect/impersonal] wedding was amusing.
The game show contestant's [incorrect/indirect] answer caused her to lose the match.
Dan criticized Joel's [inaccurate/inadequate] depiction of the story's plot.
Charlie wondered why Laurie had been [unkind/unhappy] at brunch that morning.
Jack and Fran were trying to [recall/relive] the evening of their first date.
After his stock dropped, Martin decided to [rethink/review] his investment choices.
Jim offered to [repay/return] the money he had borrowed by next week.
Paula tried to [rewrite/remove] the confusing portions of her essay.
The competitors were determined to [disclose/discover] the company's new formula.
Kim could not hide her [disgrace/dislike] when she saw the man who stood her up.
Stacey's responses were often [incorrect/indirect] because she did not do her homework.
Due to its [inefficient/inadequate] resources, the small country was quite poor.
The scientist's finding was [uncommon/uncertain] so he decided to run more tests.
Jen thought Greg's form letter was [improper/impersonal] after their 25 year friendship.
On the witness stand, the victim was forced to [relive/review] the events for the jury.
Molly's teacher asked her to [retell/return] the children's story to the class.
Ted planned to [rewrite/reform] the organization's objectives.
The construction company was hired to [rebuild/replace] the restaurant after the fire.

APPENDIX B
Experiment 2 Stimuli

Short Root Frequency Set [Low/High Frequency]

Martha wondered if this was the right pump to [refuel/refill] her car with unleaded.
The witness had to [recite/relive] the events of that awful night.
Lyn wanted to [unfold/unpack] her garment bag so her suit wouldn't wrinkle.
Upon his arrival, Nathan began to [unwrap/unpack] the gifts he'd brought.
The fruit was [unripe/unhurt] because the cellar was cold.
Hugh asked Jake to [unfold/unroll] the dirty sleeping bag out on the porch.
Before Lucy was able to [unhook/unlock] the rusty latch, she had to oil it thoroughly.
Tim tried to [unhook/unplug] the cord that was behind the couch, but it was stuck.
The town thought it was time to [retire/reopen] the old landfill.
Ira suggested they [retire/reopen] the issue.
The accountant intended to [defraud/debrief] his client, but their meeting was canceled.
When Tina's investor tried to [defraud/degrade] her, she moved her business.
The archeologist struggled to [unscrew/unearth] the container to recover its contents.
The hypnotist intended to [imprint/implant] the powerful image in Troy's mind.

Long Root Frequency Set [Low/High Frequency]

Shelley hoped to [disguise/disgrace] her famous brother at the party.
The politician's [indiscriminate/inappropriate] name calling cost him the election.
Karen found her apartment in [disarray/disorder] after the burglary.
The principal decided to [reassign/recreate] the senior class project next year.
Ally wanted to [reassign/redirect] her assistant to a new project.
The coach liked to [reaffirm/reassure] his team by showing confidence in them.
The English professor had a tendency to [misquote/misjudge] German philosophers.
George made sure to [prerinse/pretreat] the lipstick stains on his shirt.
People are annoyed by Carl's [immature/improper] behavior during his mid-life crisis.
It was Robin's responsibility to [rearrange/reprocess] the employees' work schedules.
The contestant's answers seemed [insincere/incorrect] and the host didn't notice.
The company asked the laboratory to [redeliver/redevelop] the new formula.
The governor asked the school boards to [recertify/redevelop] their math programs.
The practice appeared to be [unorthodox/unfamiliar] to the group of visitors.

Short Word Frequency Set [Low/High Frequency]

In the story, the brave hero had to [defang/defend] the lion as his final test.
The workers complained the management's policy was [unsafe/unfair] at the meeting.
Edward was determined to [defuse/defeat] Erin's motion in the Senate.
Jack and Fran tried to [relive/recall] the evening of their first kiss.
The legislator wanted to [refine/remove] the restrictions on gambling.
The editor demanded that Eric [reword/remove] the controversial statements.
Nicole thought it was [unwise/unfair] to play such a mean prank on Tommy.
Hal asked Joel to [unroll/unload] the banners and put them up for the party.
The waitress offered a [refill/refund] when Ned complained that the soda was flat.
Rick tried to [derail/defeat] the plan to replace the senior center with condos.
The manager asked Tammy to [disrobe/discuss] the mannequin for the new display.
Due to the blizzard, the [retrial/reunion] was delayed for two weeks.
Tracy wanted to [remodel/recover] her kitchen cabinets for a new look.
Frustrated by the [unequal/unaware] groups of children, the guide quit.
When the armoire didn't fit, Jamie's [inexact/invalid] measurement was to blame.

Long Word Frequency Set [Low/High Frequency]

The corporation attempted to [disclaim/discount] their competitor's accusation.
The scientist set out to [disprove/discount] his rival's theories one by one.
Gene's roommate seemed to have an [indecent/infinite] number of girlfriends.
Gina felt that her boss gave her an [inhumane/infinite] amount of work.
The fickle man told his girlfriend she was too [imperfect/impatient] and then he left.
Professor Woo wondered if the class would [disfavor/discover] his preferred study method.
Jeff complained for weeks after his [misconduct/misfortune] resulted in him losing his car.

APPENDIX B (Continued)

The teenager was told that his [indecent/informal] attire was unacceptable.
Sonia wanted to [discolor/discover] her sister's favorite shirt so she could ruin it.
Janet's aloof behavior was [unsocial/uncommon] for such a celebrated event.
Paul said he had not met the [inactive/indirect] member of the club yet.
Larry wanted to [displease/discharge] the salesman who had been so rude to him.
Jerry's vague response was [indefinite/inadequate] for making plans that weekend.
The young model was [immodest/immature] with regards to her looks.
Samantha felt [insecure/informal] wearing a casual dress at such a fancy wedding.