DOCUMENT RESUME

ED 302 829 CS 009 480

AUTHOR	Nagy, William; And Others
TITLE	Morphological Families in the Internal Lexicon. Technical Report No. 450.
INSTITUTION	Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	Jan 89
CONTRACT	OEG-0087-C1001
NOTE	36p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Higher Education; Language Research; *Lexicology; *Morphology (Languages); Reading Re3earch; Undergraduate Students; *Word Recognition
IDENTIFIERS	*Internal Lexicon; *Word Families

ABSTRACT

A study examined whether the morphological structure of words--that is to say, the analysis of words into prefixes, stems, and suffixes--plays a role in how words are represented in an individua'.'s internal lexicor Ninety-five students from a large midwestern university identified stem words, matched for length and individual frequency, which differed substantially in the frequency of their inflectional, derivational, and nonmorphological relatives. The frequency of inflectionally and derivationally related words significantly affected speed and accuracy of recognition of stems. However, the effects were conditioned by age of acquisition and part of speech. Taken as a whole, the results supported the concept of a word family, that is, the hypothesis that morphological relationships among words, derivational as well as inflectional, are represented in the lexicon. (Nine tables of data are included, and 23 references are attached. An appendix shows the reaction times and frequencies for stimulus words.) (Author/SR)

CENTER FOR THE STUDY OF READING

A READING RESEARCH AND EDUCATION CENTER REPORT

Technical Report No. 450

MORPHOLOGICAL FAMILIES IN THE INTERNAL LEXICON

William Nagy
Richard C. Anderson
Marlene Schommer
Judith Scott
Anne Sta'lman
University of Illinois at Urbana-Champaign

January 1989

University of Illinois at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

The work upon which this publication was based was supported in part by the Office of Educational Research and Improvement under Cooperative Agreement No. OEG 0087-C1001. The publication does not necessarily reflect the views of the agency supporting the research.

Abstract

The role of morphological relatedness in word recognition was examined using a lexical decision task. Ninety-five college students identified stem words, matched for length and individual frequency, which differed substantially in the frequency of their inflectional, derivational, and nonmorphological relatives. The frequency of inflectionally and derivationally related words significantly affected speed and accuracy of recognition of stems. However, the effects were conditioned by age of acquisition and part of speech. Taken as a whole, the results support the concept of a *word family*, that is, the hypothesis that morphological relationships among words, derivational as well as inflectional, are represented in the lexicon.

MORPHOLOGICAL FAMILIES IN THE INTERNAL LEXICON

This study asks whether the morphological structure of words--that is to say, the analysis of words into prefixes, stems, and suffixes--plays a role in how words are represented in people's internal lexicons. Understanding how words are stored in memory, and how it is that people can access the meaning of words so quickly, are questions that have attracted the attention of a number of researchers. Nonetheless, there are still important unresolved theoretical issues.

Our interest in this area is also motivated by concern for potential educational application. An important characteristic distinguishing good and poor readers is the ability of good readers to recognize words with great speed and efficiency. This difference appears to be more pronounced for longer words. Since most longer words are morphologically complex, and since there are large individual differences in knowledge of English derivational morphology (Freyd & Baron, 1982; Gleitman & Wanner, 1982) which are associated with differences in reading ability (Tyler & Nagy, 1985), deficiencies in morphological knowledge may be a cause of poor readers' difficulties with long words (Anderson & Davison, 1988).

We are, therefore, interested in determining the mechanisms by which morphological knowledge may contribute to efficient word recognition. Previous research has failed to find a computational cost associated with morphological complexity (Kintsch, 1974). In fact, research on the role of word frequency in word recognition suggests a mechanism by which knowledge of morphological relationships could facilitate word recognition.

For sophisticated readers, at least, decomposition of words into parts could facilitate processing, because the individual parts are more frequent than the whole word. It is a well-established finding that more frequent words are recognized more quickly than words lower in frequency. *Stand* and *steed* are the same length, but *stand* is recognized much more quickly.

The best measure of the frequency of a word is presumably not the frequency of that word alone, but the frequency of a *family of words* closely related in form and meaning. The word *inactivity*, for example, is a relatively low frequency word, occurring less than once in a hundred million words of school text, according to Carroll, Davies, and Richman (1971). If frequency were to operate independently of morphological relationships, accessing this word would be relatively slow. However, when related words such as *active*, *inactive*, *activity*, and *activities* are taken into account, the family frequency of *inactivity* is ten thousand times as great as the frequency of this individual member. If family frequency plays any role in word recognition, *inactivity* should be accessed much faster than its frequency as an individual word would lead one to expect.

The overarching theoretical question about word families is how they are represented in people's internal lexicons. One possibility is that all of the members of a family share the same lexical entry, organized under the stem, and that inflectional and derivational affixes are stripped off before the entry is accessed. For example, when faced with *untie*, a person could set aside *un-*, look up the entry for *tie* in his or her internal lexicon, and then compute that *untie* means to reverse the action of tying.

At the other extreme is the possibility that every word has a separate and distinct entry. There are several in-between possibilities. It could be, for instance, that regular inflections and, perhaps, semantically transparent derivatives share the entry with the stem, while more distant relatives have separate entries. Or, it could be that the lexicon contains separate, but linked, entries for different members of a word family.

The fact that frequency of usage has a strong and dependable effect on speed of word recognition can be used as a tool to make inferences about how word families are represented in the mental lexicon. If it were to turn out that the frequency of inflections and derivatives has as much impact on the recognition of a stem as the frequency of the stem itself, that would be strong evidence for a theory that says that the members of families share single entries. On the other hand, if the frequency of inflections and derivatives has no influence on recognition of the stem, that would be evidence for the theory that even closely related words have separate entries. In principle, in-between theories, as well, can be discriminated based on the pattern of the evidence.

Inferences about the structure of the subjective lexicon can also be made by examining the effect of the frequency of stems on speed and accuracy of recognition of their morphological relatives. Most previous research, in fact, has taken this approach. Research up to now has looked at words related in terms of four basic categories: prefixes, inflectional suffixes, derivational suffixes, and compounds.

With respect to prefixes, Taft (1979) found that the reaction time for words consisting of a prefix and a stem is influenced by the frequency of the stem. For example, *dissuade* and *reproach* are equal in frequency. But the stem *proach* also occurs in the relatively frequent word *approach*, whereas the stem *suade* occurs in the less frequent *persuade*. Taft found that in a lexical decision task, the reaction time for words like *reproach* was shorter than that for words like *dissuade*. This seems to indicate that for prefixed words, stem frequency, and not just individual word frequency, influences reaction time.

As for regular inflections, a variety of research indicates that the frequency of inflected forms influences speed of recognition of the stem word (see Cutler, 1983, and Taft, 1985, for reviews). That is, the reaction time for *walk* appears to depend, not just on the frequency of *walk*, but also on the frequency of *walks*, *walked*, and *walking*.

Lexical decision studies using repetition priming instead of the frequency effect as a basis for exploring morphological relationships show a similar, strong relationship between stems and their regular inflections. Stanners, Neiser, Hernon, and Hall (1979), for example, found that a regular inflection, such as *thinks*, primes the stem, *think*, as much as the stem primes itself.

Compounds appear to be accessed via their component parts (Andrews, 1986; Taft, 1985). That is to say, the frequency of the component parts plays a role in how quickly compounds are accessed.

The picture is least clear in the case of derivational suffixes such as *-ness* or *-ion* that change the part of speech of a word. Bradley (1979) found that the frequency of the stem influenced reaction times for some suffixed words--those ending in *-er*, *-ness*, and *-ment*--but not for words ending in the suffix *-ion*. This difference may relate to the type of suffix involved: *-er*, *-ness*, and *-ment* are neutral suffixes; they make few changes in the spelling or pronunciation of words they are added to, and are usually added only to stems that are words in their own right. The suffix *-ion*, on the other hand, is non-neutral; it is often associated with substantial changes in the spelling and pronunciation of a stem (e.g., *destroy/destruction*); and it is often found on stems that are not themselves words (e.g., *nation*).

Stanners, Neiser, Hernon, and Hall (1979) found that derivationally affixed words prime their stems less strongly than do the words themselves (or their regular inflections), suggesting that the derivatives constitute separate, though related, lexical entries. Fowler, Napps, and Feldman (1985), however, in experiment designed to disentangle the effects of episodic and lexical priming, found that derivatives primed their stems as strongly as did inflections or the stems themselves.

An objection that can be raised against much of the previous research is that the lexical decision task may not adequately represent normal reading. Specifically, when the stimuli are predominantly complex words, subjects may adopt a strategy of morphological decomposition not used in normal reading. Rubin, Beck, and Freeman (1979) have raised this objection to Taft's approach. In an experiment attempting to replicate another finding by Taft--that pseudoprefixed words like *uncle* are accessed more slowly than genuinely prefixed words like *unlike*--they found an effect only when the stimuli consisted entirely of prefixed and pseudoprefixed words. Hence, stems may be used to access prefixed words only when the stimulus set contains a large proportion of prefixed words. In a subsequent study, Taft (1981) claims to have answered the objections raised by Rubin et al., by using only prefixed words with bound stems, which do not constitute words in themselves without the prefix (e.g., *rejoice*), and pseudo-prefixed words, in which the initial letters are identical to English prefixes, but do not function as prefixes in these words (e.g., *prosaic*). However, one could still argue that every word in Taft's (1981) stimulus set contains a letter string that could be a prefix, and hence invites a strategy of prefix-stripping not used in normal reading.

Andrews (1986) explored the same objection in the case of derivationally suffixed words. She found that stem frequency affects the reaction time for derivatives, but only when most of the items are derivatives and compounds. Andrews concluded that subjects adopt special strategies when the stimuli are predominantly complex words.

Bradley's (1979) results may be subject to the same criticism. In one of the experiments, for example, about half of the stimuli ended in the suffix *-ness*. To decide whether the stimulus was a word, all the subject had to do was to strip off the suffix and determine if the remainder was a word. It is quite likely, then, that subjects' performance in this experiment do not reflect the strategies they would use in normal reading.

All in all, results of previous research indicate that morphological relationships among words are represented in some way in the internal lexicon. However, the evidence is stronger for inflectional relationships than for derivational relationships, and is least clear concerning the extent to which derivational relationships contribute to the effects of frequency. This last issue is the primary concern of the present study, because the effects of frequency on speed of lexical access is one way--although not the only one--in which a working knowledge of derivational morphology may contribute to skilled reading.[1]

The present study sought to extend previous research in several ways. First of all, we directly addressed the objection that in lexical decision experiments with a large proportion of morphologically complex words, subjects may adopt morphological decomposition as a special strategy, one that is not characteristic of normal reading.

Most research to date has asked what effect the frequency of a stem has on the reaction time for a related complex form. For example, does the speed with which someone recognizes a complex word, like *quietness*, depend on the frequency of the stem, *quiet*? Such studies necessarily contain a large proportion of derived words; hence, when an effect of stem frequency is found, it might be attributed to limited-scope strategies adopted by subjects for this task and these materials.

In the present study, this objection was finessed by addressing a different, but related question: Does the speed with which a person recognizes a stem, like *quiet*, depend just on the frequency of the stem, or is it influenced by the frequency of relatives, such as *quietly* and *quietness*? To answer this question, one need not include any morphologically complex words in the stimulus set; therefore, nothing sensitizes subjects to morphological relationships and there is no reason for them to employ special strategies. If the reaction time for *quiet* is influenced by the frequencies of *quietly* and *quietness*, this influence must come from the subjects' prior experience with these words.

Second, the present study examined three kinds of word relatedness, whereas previous studies usually have focused on one kind of relatedness. Included were words related by inflectional morphology, words related by derivational morphology, and words related because they share nonmorphological parts. Thus, the study aimed to provide a broad assessment of the concept of a word family.

Third, the study described in this paper examined the role that factors in addition to frequency may play in determining which types of relatedness among words influence word recognition and condition the effect of frequency. Reisner (1972; cited in Taft, 1985) claimed that stem frequency influences reaction times for suffixed words, but only suffixed words of lower frequency. Other factors known to affect reaction time, such as length, age of acquisition (e.g., Brown & Watson, 1987; Gilhooly, 1984), or number of meanings (Jastrzembski, 1981) may also condition the effect of frequency. It may be, for example, that suffixed words which are acquired relatively early are usually recognized without analysis into parts.

However, aside from Reisner (1972), there has been little attention paid to factors that might influence the extent to which morphological relatedness plays a role in word recognition. In this study, a number of variables have been taken into account: age of acquisition, part of speech, abstractness, oral language frequency, position of the stressed syllable, number of distinct meanings, relative frequency of the stem and its most frequent derivative, part of speech of the most frequent derivative, formal relationship between the stem and its most frequent derivative, semantic transparency of the relationship between the stem and its most frequent derivative, and, of course, word length.

Fourth, the present study involved a greater than typical number and variety of words, generated in preliminary computational linguistics spadework, and a larger than typical number of subjects. Most previous studies of the role of frequency in word recognition have been limited in size. This is worrisome because estimates of the true frequency of infrequent words are inherently unstable (Carroll, Davies, & Richman, 1971), and because individual people are likely to have idiosyncratic patterns of exposure to infrequent words (Nagy & Anderson, 1984). Thus, it can be argued that an ample number of words and subjects are necessary, not merely nice, when the effects of frequency are being investigated.

Method

Materials

The basic idea was to choose pairs of words that were matched for length and individual frequency, based on Carroll, Davies, and Richman (1971), but so that words related to the two members of each pair differed greatly in frequency. There were three sets of 28 pairs of words (see Appendix).

The first set consisted of pairs that differed in terms of inflectional family frequency, defined as the sum of the frequencies of all of the inflections of the target word, including the comparative and superlative degrees in the case of adjectives. Table 1 gives an example of a pair of words from this first set.

[Insert Table 1 about here.]

As can be seen in the table, *stair* and *spike* have the same frequency. However, the plural *stairs* is much more frequent than *spikes* and *spiked*. This set of words provides for what is essentially a replication of one of the experiments done by Taft (1979).

A second set of 28 pairs of words differed primarily in terms of derivational family frequency, defined as the sum of the frequencies of all derivatives of the target word. Table 2 gives an example of such a pair.

[Insert Table 2 about here.]

Slow and *loud* are identical in frequency. But there is quite a difference in frequency between *slowly* and *loudly*. Taking derivational relationships into account, the family frequency of *slow* is three times that of *loud*.

The third set of 28 word pairs differed in terms of still another type of relationship--nonmorphological overlap, defined as the sum of the frequencies of all words that include the target word as an embedded letter string, but which have no morphological relationship to the target word. Table 3 contains such a pair of words.

[Insert Table 3 about here.]

The words *fee* and *cod* are equal in frequency. However, the letter sequence *f-e-e* occurs in words that are many times more frequent, especially *feel* and *feet*. *C-o-d*, on the other hand, occurs only in words such as *code*, *coddle*, and *coda*, which are not anywhere near as frequent as *feel* and *feet*.

As an aid to generating candidate words, a computer program was written that searched the Carroll, Davies, and Richman (1971) corpus for words that contained any given stem as an orthographic substring. Human judges augmented the families produced by the computer so as to include irregular inflections and non-neutral derivatives. Then the aggregate frequency of each family was calculated. Words whose family frequencies were at least four times greater than their stem frequencies became candidate high family frequency words. Words whose family frequencies were less than one and a half times their stem frequencies became candidate low family frequency words.

For each candidate item, then, a set of words related by morphology or incidental orthographic overlap was identified. Each member of this set was then coded according to the type of relationship it bore to the stem--inflectional, derivational (including prefixation, suffixation, compounding, and irregular relationships such as that of *pride* and *proud*) and nonmorphological. Words related to the stem by nonmorphological overlap were also further categorized with respect to consistency of pronunciation, that is, into those for which pronunciation was the same (e.g., *fee/feet*), and those for which it changed (e.g., *cod/code*).

Finally, pairs of words were selected that were matched for length and stem frequency, but differed in the frequency of inflectional, derivational, or nonmorphological relatives. From among the candidates, word pairs were chosen for the three sets described above which maximized the family frequency difference of the type contrasted in the set and minimized--though did not completely eliminate--differences of the other two types. The low family frequency member of each pair was one for which there was little difference between stem and family frequency, but which matched the high frequency member in stem frequency and length. Word pairs were further chosen so that they encompassed a range of stem frequencies and lengths (from three to eight letters).

An attempt was also made to match the three sets for overall distribution of items by length and frequency, but due to the limited number of candidates fulfilling other constraints, the words in the third set were shorter and less frequent than those in the other two sets.

Four trained raters coded the final sets of target words on a number of other word properties, as follows:

Age of acquisition was rated on a 4-point scale of when people are likely to have first learned the meaning of the word: (a) preschool, (b) grades 1-6, (c) grades 7-12, and (d) after high school. In this case 10 raters were used. Means from the 10 raters were used in the analysis.

Number of syllables.

Position of stressed syllable.

Bigram frequencies. The sum of the bigram frequencies of each word was computed on the basis of Mayzner and Tresselt (1965). Bigram frequency tables were also constructed using Carroll, Davies, and Richman (1971), and sums of bigram frequencies based on these tables were also computed for each stimulus word.

Part of speech of stem. All target words were coded as either noun, verb, or adjective. When part of speech was ambiguous (e.g., *burn* could be a noun or a verb), the code reflected the part of speech of the base of the most frequent derivative. For example, the most frequent derivative of *burn* is *burner*, which, though itself a noun, is derived most directly from the verb *burn*. Thus, *burn* was coded as a verb.

Number of distinct meanings. Raters were asked to think in terms of truly distinct, unrelated meanings.

Abstract versus concrete. Raters were asked to categorize stimulus words as either abstract or concrete. Object names (e.g., *corpse*, *star*) and other highly imageable words which could be taken as nouns (e.g., *dent*, *rash*) were rated as concrete. Other words (e.g., *success*, *teach*, *guilt*) were rated as abstract.

Frequency of stem compared to its most frequent relative. Stems were coded as either (a) being the most frequent member of their families, or (b) as having at least one relative more frequent than the stem itself.

Oral language frequency. Stems were coded as (a) occurring more frequently in written than oral language, (b) occurring in both written and oral language with approximately the same frequency, or (c) occurring more frequently in oral than written language.

Part of speech of most frequent derivative. The most frequent derivative of each stem was identified and its part of speech coded. In almost all cases, there was one derivative far more frequent than any other.

Formal relationship of most frequent derivative to stem. Derivatives were classified as related to the stem via: (a) neutral prefix, for example, *non-*, (b) non-neutral prefix, for example, *con-*, *ab-*, (c) neutral suffix, for example, *-ness*, *-ful*, (d) non-neutral suffix, for example, *-ity*, *-ion*, (e) compounds, for example, *corkscrew*, *lampshade*, or (f) irregular derivational relationships, for example, *pride/proud*.

Semantic transparency of relationship of most frequent derivative to stem. The semantic relationship was coded on a 3-point scale: (a) transparent, for example, *educate/education*, (b) translucent, for example, *detect/detective*, *roost/rooster*, or (c) opaque, for example, *lard/larder*, *sandal/sandalwood*.

Procedure

Prior to the lexical decision task, a wide-range paper-and-pencil vocabulary test (French, Ekstrom, & Price, 1963) was administered to the subjects. Instructions for the lexical decision task were then read aloud as subjects followed along reading instructions displayed on the computer screens.

In the lexical decision task, the stimulus word or nonword appeared in the center of the screen of a personal computer, in lower case, 4-point IBM standard font. It appeared 1500 msec after the subject pressed the space bar to signal readiness for the next stimulus. The stimulus remained on the screen until the subject responded. Subjects were instructed to use index fingers to press a "yes" key if the stimulus was a word, and a "no" key if it was not a word. The dominant hand was always used for the "yes" response.

In order to ensure that subjects would be fixating the right region on the display when the stimulus was presented, a pointer was displayed on the screen indicating, but not masking, the position at which the stimulus would appear.

The complete stimulus set consisted of the 168 words, 168 nonwords matched for length, and 24 practice items (half nonwords). Nonwords conformed to the constraints of English spelling. Order of items was individually randomized for each subject.

Subjects were run in groups of 5 to 20 on IBM AT computers in a university computer lab. Software was written to utilize the internal clock of these computers. This program interrupted other processing when a response key was pressed. Consequently, variability in timing was not introduced by other computations being performed during the measurement of response times.

Subjects

Subjects were 109 undergraduates from a large midwestern university. Participation in the study was partial fulfillment of a course requirement. The data from 14 subjects were lost due to equipment failure.

Approach to Analysis

In the primary data analyses, within-subjects analysis of variance was used to compare subjects' performance on sets of word pairs, matched with respect to frequency and length, that differed with respect to the frequency of inflectional, derivational, or nonmorphological relatives. The dependent variables were proportion of errors and reaction time given a correct response. Proportion of errors was normalized with an arcsine transformation. Reaction times shorter than 200 msec were discarded and times greater than 5 seconds were recoded to 5 seconds. Reaction time was then normalized with a logarithmic transformation. A logarithmic transformation was also used to normalize the distributions of word frequencies and linearize their relationships with proportion correct and reaction time.

Subsidiary analyses, in which the word was the unit of analysis, were completed within the framework of the general linear model. The dependent variables in these analyses were mean log reaction time for correct responses, and proportion of incorrect responses. The logit transformation was used to normalize the distribution of error rates. The subsidiary analyses examined not just the effects of family frequency, but also the extent to which other variables such as stem frequency, age of acquisition, and length might condition the effects of family frequency. Each of the three sets of target words were analyzed separately. Analyses were also performed on the combined set.

Results

The basic results of the experiment appear in Table 4. For all three sets of target words, subjects responded significantly more quickly to those words for which the total frequency of the set of related words was greater--whether the relationships were inflectional, derivational, or purely orthographic.

[Insert Table 4 about here.]

The error data showed a slightly different pattern. For the first two sets, representing inflectional and derivational relationships, subjects made fewer errors on words with higher family frequencies. For the third set of words, however, frequency of orthographically overlapping words did not significantly influence subjects' error rates.[2]

We will now examine the results for each set of words in detail.

Inflectional Relationships

For the first set of words--pairs of words differing primarily in the frequency of their inflections, like *stair* and *spike*, subjects responded significantly more quickly to words with higher inflectional family frequencies, $F(1,94) = 39.2$, $p < .001$. They also made fewer errors on such words, $F(1,94) = 35.5$, $p < .001$.

Results of regression analyses for the first set of words with reaction time as the dependent variable, and using the word as the unit of analysis are given in Table 5. There is a significant effect of inflectional frequency; but this effect disappears if age of acquisition is included in the analysis, as can be seen in Table 6. The interaction of Inflectional Frequency x Length seen in Table 6 appeared because inflectional frequency speeded reaction times for longer stems to a greater extent than shorter stems. (In these and following tables, reduced regression models are given, in which factors not involved in significant effects have been eliminated.)

[Insert Tables 5 and 6 about here.]

A similar analysis was done using error rate (normalized using the logit transformation) as the dependent variable, and the word as the unit of analysis. In this analysis, the effect of inflectional frequency on error rate was not significant, $F(1,54) = 3.5$, $p > .05$.

Derivational Relationships

For the second set of words--pairs of words differing primarily in the frequency of their derivations, like *slow* and *loud*, subjects responded significantly more quickly to words with higher derivational family frequencies, $F(1,94) = 9.0$, $p < .01$. They also made fewer errors for such words, $F(1,94) = 42.1$, $p < .001$. These results support the hypothesis that derivational relationships among words play a role in the frequency effect.

The results of a regression analysis on the second set of words using the word as the unit of analysis are given in Table 7.

[Insert Table 7 about here.]

In this analysis, the main effect of derivational frequency was not significant. (This was the case whether or not age of acquisition was included in the analysis.) But derivational frequency interacted significantly with part of speech (coded in orthogonal contrasts); derivational frequency speeded recognition of nouns and verbs, but not adjectives.

In short, derivational relationships do make a contribution to the frequency effect, but this effect does not appear to generalize across all kinds of words.

A similar analysis was done using error rate (normalized using the logit transformation) as the dependent variable, and the word as the unit of analysis. In this analysis, the effect of derivational frequency on error rate was not significant, $F(1,54) = 1.05$, $p > .05$.

Nonmorphological Relationships

In the third set of words, nonmorphological relationships had a significant effect on reaction time when the subject was the unit of analysis, $F(1,94) = 10.6$, $p < .01$. However, the difference in error rates for subjects was not significant, $F(1,94) = 2.37$, $p > .05$.

In regression analyses using the word as the unit of analysis, the effect of nonmorphological relationships on reaction time was not significant. Table 8 displays this analysis.

[Insert Table 8 about here.]

Among the relatives of words in the third set, there were also some true derivatives, for example, *codfish*. As is evident from the analysis presented in Table 8, these true derivatives, although much lower in frequency than the nonmorphologically related words, had a highly significant effect on reaction time.

The Derivational Frequency x Part of Speech interaction found in the second set of target words was not found for this third set of words. However, the third set turned out to be composed almost entirely of nouns. Nouns in the second set, along with verbs, were influenced by derivational frequency. Based on this, one would expect to find a significant effect of derivational frequency with the third word set, and the two analyses yield consistent findings after all.

The effect of the derivational relationships in this third set of words provides an explanation for the significant effect of nonmorphological frequency in the analysis using the subject as the unit of analysis. In additional analyses using the subject as the unit of analysis, the pairs of words in the third set were divided into three groups: Those for which there were no derivationally related words that might contribute to a difference in reaction times, those for which the effect of derivational frequencies would coincide with any effect of nonmorphological relationships, and those for which the effect of derivational frequencies would be in the direction opposite to that of nonmorphological relationships.

For the first group of words--those for which derivational frequency played no role--the effect of nonmorphological relationships was small but significant, $F(1,94) = 4.85$, $p < .05$; words with higher-frequency nonmorphological relationships were recognized a little more quickly. For the second group, in which derivational and nonmorphological relationships coincided in the direction of their effects, there was a strong effect, $F(1,94) = 22.0$, $p < .001$. For the third pair of words, the difference was significant, $F(1,94) = 6.48$, $p < .05$, and the difference was in the opposite direction. That is, it was the frequency of derivational relationships, not nonmorphological relationships, that predicted which words were recognized most quickly. It should be noted that this was the case even though the magnitude of the frequencies of nonmorphologically related words was much greater than that of the frequencies of the derivationally related words.

An analysis was also performed using error rate (normalized using the logit transformation) as the dependent variable, and the word as the unit of analysis. In this analysis, the effect of derivational frequency on error rate was significant, $F(1,54) = 7.14$, $p < .01$. Nonmorphological relationships did not significantly affect error rate.

Analysis of Combined Sets

Table 9 presents the results of an analysis of all the target words together. When the entire stimulus set is considered, the effect of derivational frequency is still significant, as is the Derivational Frequency x Part of Speech interaction.

[Insert Table 9 about here.]

Other Word Properties

In the subsidiary analyses of the second and third sets of stimulus words, other word properties were investigated to see if they might condition the effect of derivational frequency on reaction time. Investigated were properties of the stimulus words, such as stem frequency, age of acquisition, length in letters and syllables, abstractness or concreteness, position of the stressed syllables, number of distinct meanings, relative magnitude of oral and written frequencies, and whether the most frequent derivative was more or less frequent than the stem; and properties of the most frequent derivative of the stimulus word, such as type of derivational relationship (prefixation, suffixation or compounding), part of speech of the derivative, semantic transparency of the derivational relationship, and whether the prefixes and suffixes were neutral or non-neutral. None of the two-way interactions of these variables with derivational frequency were found to be significant. Nor, except as already noted, were any of the main effects of these variables significant. In particular, neither measure of bigram frequency made a significant contribution to predicting reaction time.

Ability

Subject verbal ability, as represented by the score from the wide-range vocabulary test, did not interact with any measure of family frequency in the analyses using the subject as the unit of analysis. We interpret this to reflect restriction in range--that is, the derivational relationships represented in our materials were in general known to all the college undergraduate subjects.

Discussion

The primary purpose of this study was to determine the effect of relationships among words on word recognition, specifically, the extent to which speed and accuracy in identifying stems are influenced by the aggregate frequency of related words, and for what types of relationships among words this influence may hold. At issue is how related words are represented in the internal lexicon. At one extreme, it might be supposed that every word has a completely separate entry. At the other extreme, it could be theorized that morphologically complex words are accessed always and only through their component stem morphemes.

The significant effect of inflectional and derivational family frequency on the speed and accuracy with which stems are identified clearly rules out any theory which says that the lexicon consists of totally unconnected entries. Our results rule out the other extreme position as well. The theory that a derived or inflected word is accessed via its stem can be evaluated by looking at the relative size of the influence on reaction time of stem frequency, on the one hand, and inflectional and derivational family frequency, on the other. Our reasoning is that if the lexical entry for the word *joins*, for example, is accessed via the entry for *join*, then every encounter with *joins* will count exactly the same as an encounter with *join*. Therefore, if the theory is correct, the frequency of *joins* will contribute as much to the priority (or strength, position in push down stack, activation threshold, etc.) attached to the *join* entry, and to speed of access, as does the frequency of the stem *join* itself.

In fact, our data show that the effect of the frequency of morphologically related words on the reaction times for stems, though significant, is not as strong as the effect of the frequency of the stems themselves. Regression analyses were performed in which the effects of inflectional, derivational, and stem frequency on reaction time were examined independently, to give each type of frequency measure full credit for any shared variance. The resulting regression equations were then used to calculate the change in reaction time resulting from a 10-fold gain in frequency, for instance, from 10 times in a million words of text to 100 times in a million words of text.

A change in stem frequency of this size leads to a 63 msec decrease in reaction time. The same size change in inflectional family frequency leads to a 19 msec decrease, and for derivational family frequency, the decrease is 20 msec. In other words, an encounter with the words *decided* and *decision* affects the speed of a person's future response to the word *decide*, but not as strongly as does an encounter with the word *decide* itself.[3]

Nonmorphological relationships have the opposite effect. A ten-fold increase in the nonmorphological relatives results in a 7 msec *increase* in reaction time. If one looks at those words for which there is orthographic overlap, but a change in pronunciation (e.g., the relationships *code* bears to *cod*) the effect is to slow down word recognition even more; a ten-fold increase in the frequency of such nonmorphological relatives leads to a 23 msec increase in reaction time.

The theory to explain the foregoing facts is that the lexicon is organized so that the entries for related words are linked. Accessing any one of the entries causes partial activation (cf. Stanners, Neiser, Hernon, & Hall, 1979) of related entries. Thus, there are subentries under the stem or linked main entries for *join*, *joins*, and *joint*, and for *decide*, *decided*, and *decision*, and accessing one of the words in either family partially activates other family members.

Note the equivalence in size of the effects of inflectional and derivational frequency on reaction time for the stem. Since linguistic theory and previous findings leave little doubt that inflectional relationships must be represented in some way in the internal lexicon, this is an additional reason for believing that, despite some complications in our results, derivational relationships are represented in the lexicon as well.

It is important to determine not just *whether* relationships among words are represented in the internal lexicon, but also *which* relationships. Our results suggest that it is morphological relationships, and not simply overlap among word parts, that are important. To be sure, our results were complicated by the fact that the frequency of nonmorphologically related words happened to be correlated with the frequency of derivational relatives. When this confound was statistically discounted, however, the apparent association between frequency of nonmorphologically related words and reaction time nearly disappeared. Moreover, nonmorphological relationships were never close to being significant in analyses in which the word was the unit of analysis. Indeed, in these analyses, the sign of the regression coefficients pointed in the wrong direction when inflectional and derivational family frequency were in the equation. The fact that bigram frequencies were not significantly related to reaction times also supports the conclusion that it is morphological, rather than purely orthographic, relationships that play a role in speeding word recognition. Thus, our findings are consistent with those of Murrell and Morton (1974), who found effects of inflectional relationships among words, but no effects of non-morphological relationships which involved the same degree of orthographic similarity as the inflectional relationships.

Results of other studies (e.g., Taft, 1979) would lead to the expectation that inflectional frequency would strongly influence reaction times for stems. Hence, it may appear puzzling that the effect of inflectional frequency on reaction time for stems disappeared when age of acquisition was included in the analysis. Our explanation is that it is likely that inflectional frequency is naturally confounded with age of acquisition. For example, in a pair of words like *stair* and *spike*, not only does the infrequent stem *stair* have a very frequent plural, *stairs*, but *stairs* is also acquired much earlier. In addition, age-of-acquisition ratings probably include a component of subjective frequency, and subjective frequency in turn would be expected to include inflectional family frequency.

Results from other types of experiments make it clear that inflectional frequency should be expected to influence reaction time. Therefore, we interpret our results as showing, not that inflectional frequency fails to influence reaction time, but that it is difficult to distinguish the effects of inflectional frequency and age of acquisition in a lexical decision experiment of this sort. Nevertheless, the results do suggest that the interpretation of previous research, which did not take age of acquisition into account, is not as straightforward as had been thought.

The primary question addressed in this study is about the status of *derivational* relationships. Previous research does not give an unequivocal answer to this question. Bradley (1979) found effects of stem frequency on the reaction time of derivatives for some suffixes but not others. Andrews (1986) found stem frequency affected reaction time to suffixed derivatives, but only when the stimulus list contained large numbers of complex words. Andrews did find effects of stem frequency for compounds, but these results and those of Bradley could also be accounted for in terms of the idea that subjects adopt special strategies when confronted with mostly complex words.

The present experiment constituted a stringent test of the claim that the entries of stems and derivatives are interconnected, because the stimulus set contained no morphologically complex words. Hence, there was no reason for subjects to bring into play any strategy that entailed focusing on morphological relationships.[4]

Because all the stimuli in the present experiment were stems, different types of derivational relationship are not directly represented. However, for almost all the stimulus words with a high derivational family frequency, there was a single derivative much more frequent than any other, which accounted for most of the derivational family frequency of that stem.

In the second set of stimulus words, all but one of those 28 stem words with high derivational frequency had a suffixed word as its most frequent derivative. Of the 27 suffixed words, 17 had neutral suffixes, and 10 non-neutral suffixes. It is noteworthy that neutrality of suffix had absolutely no influence on the extent to which derivational frequency affected reaction time for the stem ($F < 1$). This result is inconsistent with Bradley's (1979) finding that stem frequency decreases reaction time for derivatives terminating with the neutral suffixes *-er*, *-ess* and *-ment*, but not derivatives ending in the non-neutral suffix *-ion*.

Half of the 56 words in the third set of stimulus words had derivatives that appeared in Carroll, Davies, and Richman (1971). The most frequent derivative of 14 of these was a compound, and for 13, the most frequent derivative was suffixed. (All but one of the suffixed derivatives had a neutral suffix.) However, whether the most frequent relative was a compound or suffixed derivative had no influence on the extent to which derivational frequency affected reaction time for the stem ($F < 1$).

We were surprised by the fact that whether a derivative was neutral or non-neutral did not seem to matter, but we were even more surprised by the Derivational Frequency x Part of Speech interaction, which appeared because derivational family frequency influenced reaction time for verbs and nouns but not adjectives. We do not know of any similar finding in the literature, nor of any psycholinguistic explanation for why derivational family frequency should not affect reaction time for adjectives. As it happens, the most frequent derivatives of the adjectives employed in this study were all adverbs formed with the highly productive suffix *-ly*. One would have expected these words to show a stronger effect of derivational frequency, if anything.

The Derivational Frequency x Part of Speech interaction tells us what was already evident from the fact that derivational relationships showed a significant effect on reaction time in the by-the-subject analysis, but not in the by-the-word analysis: The frequencies of derived words influence speed of recognition for some stems, and not for others.

That part of speech is the critical factor seems implausible; the effect might be the result of some uncontrolled property of the adjectives or their relatives. Before lightly dismissing the effect of part of speech, though, it ought to be stressed again that in this study numerous properties that might condition the effects of derivational frequency were examined: Properties of the stem, such as stem frequency, age of acquisition, length in letters and syllables, abstractness or concreteness, position of the stressed syllable, number of distinct meanings, relative magnitude of oral and written frequencies, and whether the most frequent derivative was more or less frequent than the stem; and properties of the most frequent derivative of the stem, such as type of derivational relationship, semantic transparency of the derivational relationship, and whether the prefixes and suffixes were neutral or non-neutral. But despite the rather exhaustive (and exhausting) search, no confounding variable was discovered which could explain the failure of derivational family frequency to affect speed of response to adjectives.

To recapitulate, the present study supports the concept of a *word family*. The results suggest that morphologically related words are grouped together under the same entry in the internal lexicon or, perhaps, in linked main entries. The most newsworthy result was the fact that derivational family frequency had the same effect on stem reaction time as inflectional family frequency. Passing over the complication involving adjectives, this result makes the *prima facie* case that derivatives and inflections are represented in a comparable manner in the lexicon.

Andrews' (1986) study showed that morphological decomposition is a strategy that skilled readers *can* adopt, given special task demands. However, in Andrews' study, this strategy was not adopted unless the stimulus materials contained a high proportion of complex words. This would imply that during normal reading, skilled readers' knowledge of derivational morphology does not play a role in word recognition.

Our results indicate the opposite. Because we were looking at the effect of derivational frequencies on reaction times for stems, our results reflect not whatever morphological decomposition may take place during a lexical decision task, but rather, the cumulative results of morphological decomposition during the subjects' years of language use. Thus, the present study strongly suggests that knowledge of morphology plays a role in word recognition during normal reading.

References

Anderson, R. C., & Davison, A. (1988). Conceptual and empirical bases of readability formulas. In G. Green & A. Davison (Eds.), *Linguistic complexity and text comprehension* (pp. 23-54). Hillsdale, NJ: Erlbaum.

Andrews, S. (1986). Morphological influences on lexical access: Lexical or nonlexical effects? *Journal of Memory and Language, 25*, 726-740.

Bradley, D. C. (1979). Lexical representation of derivational relation. In M. Aronoff & M. L. Kean (Eds.), *Juncture* (pp. 37-55). Cambridge, MA: MIT Press.

Brown, G. D. A., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory and Cognition, 15*, 208-216.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage word frequency book*. Boston: Houghton Mifflin.

Cutler, A. (1983). Lexical complexity and sentence processing. In G. B. Flores d'Arcais & R. J. Jarvella (Eds.), *The process of language understanding* (pp. 43-79). New York: Wiley.

French, J. W., Ekstrom, R. B., & Price, L. A. (1963). *Kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Service.

Freyd, P., & Baron, J. (1982). Individual differences in acquisition of derivational morphology. *Journal of Verbal Learning and Verbal Behavior, 21*, 310-332.

Gilhooly, K. J. (1984). Word age-of-acquisition and residence time in lexical memory as factors in word naming. *Current Psychological Research and Reviews, 3*, 24-31.

Gleitman, L., & Wanner, E. (1982). Language acquisition: The state of the state of the art. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3-48). Cambridge: Cambridge University Press.

Jastrzembski, J. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology, 13*, 278-305.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.

Mayzner, M. S., & Tresselt, M. E. (1965). Tables of single letter and digram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements, 1*(2), 13-32.

Murrell, G. A., & Morton, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology, 102*, 963-968.

Nagy, W., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*, 304-330.

Reisner, P. (1972). *Storage and retrieval of polymorphemic words in the internal lexicon*. Unpublished doctoral thesis, Lehigh University.

Rubin, G. S., Becker, C. A., & Freeman, R. H. (1979). Morphological structure and its effect on visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 18*, 757-767.

Stanners, R., Neiser, J., Hernon, W., & Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior, 18*, 399-412.

Stanners, R., Neiser, J., & Painton, S. (1979). Memory representations for prefixed words. *Journal of Verbal Learning and Verbal Behavior, 18*, 733-743.

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

Taft, M. (1981). Prefix stripping revisited. *Journal of Verbal Learning and Verbal Behavior, 20*, 289-297.

Taft, M. (1985). The decoding of words in lexical access: A review of the morphographic approach. In D. Bresnan, T. G. Waller, & G. E. Mackinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 83-123). New York: Academic Press.

Tyler, A., & Nagy, W. (1985). *The role of derivational morphology in sentence comprehension* (Tech. Rep. No. 357). Urbana: University of Illinois, Center for the Study of Reading.

Author Notes

The authors would like to thank David Zola and Jerry Wilson for providing software for the lexical decision task, and for other computational help.

Footnotes

[1]Knowledge of derivational morphology is likely to play an even more crucial role in dealing with new words. More than 60% of the new words that readers encounter have relatively transparent morphological structures--that is, they can be broken down into parts, at least some of which are themselves words, and the meanings of these parts give sufficient information to make a good guess at the meaning of the whole word (Nagy & Anderson, 1984).

[2]As can be seen in Table 4, the stimuli in the third set of words had longer reaction times and higher error rates than those in the first two sets. This is probably because, as explained in the Methods section, a higher priority had to be placed on matching length and frequency in pairs within sets than matching overall length and frequency between sets. The stimulus words in the third set were on the average shorter but less frequent than those in the other two sets.

[3]As a methodological matter, investigators may wonder whether, when examining word frequency as one of the variables influencing speed of recognition, it is enough to simply use the frequency of the word itself--or should one take the trouble to include the frequencies of inflections and derivatives.

Although this study shows that the frequency of morphological relatives affects word recognition, it does not follow that family frequencies account for much variation in reaction times. Stem frequencies are correlated with inflectional (r = .35) and derivational (r = .47) frequency, and the unique contribution of family frequencies is relatively small. As one adopts increasingly inclusive definitions of "word family," the gain in accuracy at predicting reaction times increases only slightly. Taking all our stimulus words together, the correlation between the log of reaction time and the individual word frequency is -.658. If one includes regular inflections in the computation of family frequency, the correlation between frequency and reaction time increases to -.695. If one includes derivational relationships as well, the correlation becomes -.699, a very small gain. (If one goes on to include orthographically but not morphologically related words, the correlation drops to -.500.) In other words, if one is interested in accuracy at predicting reaction times, it might be worth it to take inflectional relationships into account, but adding up derivational frequencies is simply not worthwhile.

On the other hand, the results indicate that as a practical, or methodological matter, age of acquisition is a factor well worth taking into consideration. From the simple rs in the tables, it is apparent that the age of acquisition ratings are about as good a predictor of reaction time as is word frequency--better, if one takes the curvilinearity of the relationship between age of acquisition and reaction time into account.

Obviously, there is a lot of overlap between frequency and age of acquisition; their correlation is -.61. But age of acquisition ratings make a significant independent contribution to reaction times even when entered after frequency. It is especially surprising that a relatively low-cost measure--10 raters rating on a 4-point scale--had as much predictive power as frequencies based on a five million word corpus.

[4]There is one reason why we may have found derivational relationships influencing the frequency effect where Andrews (1986) did not. This has to do with possible asymmetries in the priming or activation relationships among related entries. Derivatives contain their stems, but not vice-versa: When one accesses the word *painful*, to the extent that the entries are related, one must also activate the word *pain* to some extent. On the other hand, it may be possible to access the word *pain* without activating *painful*. In fact, Stanners, Neiser, and Painton (1979) found that *unhappy* fully primed *happy*, whereas *happy* did not fully prime *unhappy*.

Table 1

Sample Word Pair Differing Primarily in Inflectional Frequency

Relationship Category	Word	Frequency	Word	Frequency*
Target Word	stair	2.4	spike	2.4
Inflections	stairs	29.4	spikes spiked	3.1 0.9 --- 4.0
Derivatives	stairway staircase	5.5 2.6 --- 8.1	spiky	0.3
Nonmorphological Relationships	---		---	

*Word frequency expressed in trequency per million words of text

Table 2

Sample Word Pair Differing Primarily in Derivational Frequency

Relationship Category	Word	Frequency	Word	Frequency
Target Word	slow	68.7	loud	69.1
Inflections	slowed	11.4	louder	18.5
	slowing	4.0	loudest	1.5
	slows	3.0		----
	slower	8.5		20.0
	slowest	0.6		

		27.5		
Derivatives	slowly	205.0	loudly	21.5
			loudness	1.6
			loudspeaker	1.6

				24.7
Nonmorphological Relationships	---		---	

Table 3

Sample Word Pair Differing Primarily in Frequency of Nonmorphological Relatives

Relationship Category	Word	Frequency	Word	Frequency
Target Word	fee	3.1	cod	3.3
Inflections	fees	1.5	----	
Derivatives	----		codfish	1.5
Nonmorphological Relationships	feeble	2.7	code	21.5
	feed	65.6	codes	1.2
	feeder	3.2		----
	feeding	20.9		22.7
	feeds	7.9		
	feel	226.8		
	feelers	4.3		
	feeling	87.8		
	feelings	37.1		
	feels	30.4		
	feet	463.3		

		950.0		

Table 4

Reaction Times (in msec) and Error Rates (in Percents)

	RTs	Errors
Word pairs differing primarily in inflectional frequency		
High inflectional frequency	720	3.7
Low inflectional frequency	759	7.6
Word pairs differing primarily in derivational frequency		
High derivational frequency	730	2.7
Low derivational frequency	741	5.6
Word pairs differing primarily in frequency of nonmorphologically related orthographic overlap		
High group frequency	781	12.8
Low group frequency	795	14.4

Note. Figures have been backtransformed from mean log reaction time and mean arcsine proportion errors, respectively.

Table 5

Multiple Regression Analysis for Words in Pairs Differing Primarily in Inflectional Frequency (Without Age of Acquisition)

Variable	F	R^2	R^2 Change	Simple r
Stem Frequency	24.4*	.33	.33	-.58
Inflectional Frequency	9.1*	.46	.12	-.58

*Critical value (1,55) = 4.02, $p < .05$

Table 6

Multiple Regression Analysis for Words in Pairs Differing Primarily in Inflectional Frequencies (With Age of Acquisition)

Variable	F	R^2	R^2 Change	Simple r
Age of Acquisition	36.7*	.38	.38	.62
Age of Acq. Squared	7.9*	.47	.08	.66
Stem Frequency	9.9*	.57	.10	-.58
Inflectional Frequency	1.0	.58	.01	-.58

*Critical value (1,55) = 4.02, $p < .05$

Table 7

Multiple Regression Analysis for Words in Pairs Differing Primarily in Derivational Frequency

Variable	F	R^2	R^2 Change	Simple r
Stem Frequency	58.9*	.54	.54	-.74
Age of Acquisition	7.3*	.61	.07	.69
Age of Acq. Squared	5.6*	.66	.05	.72
Derivational Frequency	0.2	.66	.00	-.38
Part of Speech	0.7	.67	.00	----
Derivational Frequency x Part of Speech	5.3*	.72	.05	----

*Critical value (1,55) = 4.02, $p < .05$

Table 8

Multiple Regression Analysis for Words in Pairs Differing Primarily in Nonmorphological Frequency

Variable	F	R^2	R^2 Change	Simple r
Age of Acquisition	37.6*	.37	.37	.61
Length in Letters	5.9*	.43	.06	-.13
Derivational Frequency	17.6*	.60	.17	-.58
Nonmorphological Frequency	0.3	.61	.00	-.07

*Critical value (1,55) = 4.02, $p < .05$

Table 9

Multiple Regression Analysis for All Words

Variable	F	R^2	R^2 Change	Simple r
Stem Frequency	160.0*	.43	.43	-.66
Age of Acquisition	34.9*	.53	.09	.65
Age of Acq. Squared	8.4*	.55	.02	.66
Length in Letters	5.9*	.56	.02	-.18
Derivational Frequency	10.3*	.59	.03	-.51
Part of Speech	1.7	.60	.01	----
Derivational Frequency x Part of Speech	7.0*	.62	.02	----

*Critical value (1,167) = 3.91, $p < .05$

Appendix

Reaction Times and Frequencies for Stimulus Words

(a) Words in pairs differing primarily in inflectional frequencies

Word	RT	Frequencies			
		Stem	Inflectional	Derivational	Non Morphological
flop	731	0.864	3.044	0.630	0.000
loam	857	0.845	0.000	0.011	0.000
wing	700	26.310	88.220	2.595	0.016
dull	670	25.274	0.770	1.235	0.395
burn	662	30.621	112.938	3.524	1.900
lamp	644	30.752	10.708	1.044	0.153
glove	667	6.302	8.435	0.000	0.011
grief	620	6.324	0.335	3.967	0.000
stair	645	2.515	29.367	7.916	0.000
spike	650	2.508	4.044	0.313	0.000
gland	703	2.864	8.609	0.918	0.000
quail	705	2.891	0.054	0.000	0.000
pebble	696	5.678	9.145	0.136	0.000
cradle	689	5.991	1.381	0.096	0.000
mitten	777	0.237	4.811	1.681	0.000
martyr	818	0.237	0.238	0.015	0.000
muscle	660	18.758	51.447	0.079	0.000
statue	693	20.269	3.437	2.052	0.000
bruise	688	0.570	4.579	0.015	0.000
chaste	753	0.578	0.000	0.074	0.000
termite	743	1.116	3.194	0.000	0.000
treason	750	1.121	0.000	0.011	0.000
glitter	662	1.211	8.911	0.000	0.000
methane	837	1.224	0.000	0.021	0.000
whisper	649	10.867	43.923	0.071	0.000
chimney	764	11.940	4.122	0.102	0.000
scatter	664	3.940	33.218	0.119	0.000
crimson	750	3.918	0.117	0.069	0.000
bean	702	12.483	31.506	0.911	0.004
fist	684	11.076	5.089	0.023	0.021
tuck	788	1.241	9.647	0.014	3.060
shun	821	1.141	0.262	0.012	0.170
wave	688	54.274	148.146	11.038	0.068
bowl	690	52.580	14.862	1.341	0.057
choke	662	1.093	6.677	0.021	0.000
latch	677	1.117	0.091	0.000	0.000
crush	640	3.132	16.427	0.023	0.000
slang	686	3.114	0.000	0.000	0.000
troop	672	3.987	26.762	1.275	0.000
spout	734	3.326	0.961	0.000	0.000
friend	638	155.041	195.239	69.695	0.000
spring	618	156.119	18.881	5.843	1.828

(a) continued

Word	RT	Frequencies			
		Stem	Inflectional	Derivational	Non Morphological
parent	620	8.628	80.348	1.076	15.478
carpet	642	8.743	3.235	0.180	0.000
sandal	831	0.534	3.002	0.000	0.549
mirage	800	0.502	0.119	0.000	0.000
mingle	727	0.552	2.885	0.000	0.000
gospel	699	0.746	0.011	0.000	0.000
attach	663	5.964	37.361	1.634	0.000
barley	726	6.903	0.000	0.054	0.000
grumble	735	0.686	5.424	0.000	0.000
trundle	816	0.806	0.146	0.000	0.000
stumble	670	1.949	9.525	0.116	0.000
stubble	703	1.726	0.113	3.958	0.070
install	660	2.022	8.434	4.283	0.000
sulphur	874	2.125	0.000	0.059	0.000

(b) Words in pairs differing primarily in derivational frequencies

Word	RT	Stem	Inflectional	Derivational	Non Morphological
		Frequencies			
reap	773	0.927	0.237	2.703	1.991
skew	837	0.928	0.000	0.290	0.000
roost	753	1.661	0.777	9.987	0.000
yacht	730	1.955	0.362	0.000	0.000
weigh	644	28.872	49.204	115.240	0.000
false	638	28.293	0.000	0.751	0.256
teach	599	35.323	21.685	183.095	0.000
globe	640	33.912	1.611	3.059	0.012
quick	597	70.048	6.442	205.321	0.274
bread	635	77.095	0.137	1.359	1.355
dread	736	4.325	3.641	8.687	0.008
wierd	640	4.572	0.000	0.012	0.000
detect	771	4.042	3.891	14.412	0.000
pardon	712	4.016	0.463	0.022	0.000
cosmos	826	0.342	0.000	2.144	0.000
phylum	1085	0.335	0.020	0.000	0.255
sudden	648	38.281	0.000	168.060	0.000
effort	700	39.060	20.491	0.829	0.000
beauty	610	60.441	1.937	192.147	0.000
plenty	618	55.702	0.011	9.097	0.000
govern	735	4.871	9.424	191.505	0.000
random	635	4.971	0.000	0.000	0.000
explode	713	2.018	7.239	22.855	0.000
lacquer	893	2.038	0.237	0.023	0.000
success	632	43.750	1.900	97.549	0.000
balance	647	44.129	13.934	0.058	0.019
probable	810	4.481	0.000	215.260	0.000
adequate	758	4.576	0.000	3.261	0.000
slow	632	68.991	27.432	208.899	0.000
loud	628	69.119	20.007	25.053	0.015
chat	700	1.258	0.996	14.351	1.246
plop	832	1.224	0.487	0.041	0.000
pride	646	24.144	0.195	82.346	0.046
thumb	634	25.584	2.380	1.435	0.008
exact	665	40.034	0.296	120.466	0.000
sorry	648	39.704	0.000	11.843	0.000
guilt	663	1.365	0.000	7.140	0.000
steed	799	1.496	0.114	0.000	0.000
grace	655	11.989	0.608	16.555	0.187
trout	672	7.239	0.000	0.000	0.000
fright	684	9.803	0.011	63.925	0.000
priest	700	9.776	9.563	0.401	0.315
nature	638	83.010	5.341	171.358	0.000
cotton	687	80.731	0.150	5.319	0.000
critic	802	1.472	4.812	16.619	0.000

(b) continued

Word	RT	Stem	Inflectional	Derivational	Non Morphological
		Frequencies			
faucet	715	1.475	0.537	0.000	0.000
prompt	669	1.472	0.973	9.141	0.000
corpse	743	1.823	0.402	0.021	0.000
relate	672	5.581	66.389	121.207	0.000
cradle	698	5.991	1.381	0.096	0.000
educate	690	1.463	13.205	56.592	0.000
qualify	656	1.580	2.644	2.711	0.000
possess	797	5.482	9.147	31.205	0.000
harpoon	800	5.787	0.387	0.117	0.000
fortune	611	17.934	3.921	39.871	0.104
slender	674	14.338	0.291	0.000	0.000

(c) Words in pairs differing primarily in non-morphological frequencies

Word	RT	Frequencies: Stem	Inflectional	Derivational	Non Morphological
din	880	2.502	0.000	0.000	132.401
gem	810	2.674	2.376	0.012	1.489
fee	770	3.137	1.498	0.000	951.907
cod	770	5.462	0.000	1.511	28.810
bask	756	0.340	0.845	0.000	77.706
lard	701	0.340	0.003	0.510	0.011
carp	748	1.495	0.028	0.012	27.868
snip	803	1.481	1.368	0.000	0.210
dram	986	0.022	0.000	0.046	35.650
quay	861	0.023	0.169	0.000	0.000
doll	710	23.882	14.980	0.919	116.024
shed	712	23.164	3.210	0.000	0.011
bran	820	0.829	0.000	0.000	123.653
bawl	789	0.820	0.715	0.000	0.000
wand	711	1.304	0.113	0.000	38.396
doom	676	1.299	1.626	0.183	0.000
beet	770	0.474	6.292	0.039	14.866
hoax	805	0.469	0.000	0.000	0.000
germ	727	2.939	9.815	1.712	126.112
fawn	700	2.975	0.524	0.011	0.000
dill	823	1.049	0.032	0.000	3.302
kilt	861	0.622	0.019	0.000	0.091
mush	744	2.326	0.000	0.914	9.984
hemp	844	2.315	0.000	0.015	0.012
pain	645	31.496	5.435	10.366	115.036
silk	655	32.283	2.915	9.236	0.000
whisk	694	1.333	1.364	0.000	11.828
knoll	811	1.369	0.117	0.000	0.000
cot	803	2.391	0.020	0.000	101.818
pry	870	2.413	0.668	0.000	0.134
pea	677	5.568	7.685	20.662	200.760
ash	663	5.764	13.236	0.342	20.338
chap	816	1.875	1.658	0.000	81.836
wren	863	2.045	0.857	0.000	3.781
dent	698	1.063	0.701	0.000	13.350
rump	784	0.999	0.407	0.000	0.604
barb	762	0.295	3.792	0.718	22.990
muck	820	0.294	0.000	0.057	0.000
mutt	814	0.442	0.000	0.000	14.868
dint	777	0.113	0.000	0.000	0.000
spar	841	1.781	0.924	0.000	67.034
glee	735	1.780	0.000	1.302	0.011
tick	735	4.065	3.692	0.578	41.643
dike	814	3.785	3.333	0.000	0.000
star	649	89.169	121.236	11.591	669.737

(c) continued

Word	RT	Frequencies			
		Stem	Inflectional	Derivational	Non Morphological
noun	724	65.672	43.511	0.173	0.000
exam	638	0.525	0.451	74.366	449.595
levy	783	0.589	0.496	0.000	0.230
volt	766	0.604	0.864	1.844	3.310
null	841	0.601	0.000	0.132	0.046
wick	771	1.032	0.261	0.000	12.252
rash	687	1.040	0.011	0.132	0.000
sigh	756	7.621	27.086	0.046	119.922
cork	705	8.464	0.160	0.514	0.091
poise	734	1.020	2.585	0.000	26.209
niece	712	1.058	0.025	0.000	0.000

Note:
Frequencies are expressed in terms of frequency per million words of text. RTs are expressed in milliseconds, backtransformed from mean log reaction time.