Obligatory decomposition in reading prefixed words

Marcus Taft and Sam Ardasinski
School of Psychology, University of New South Wales, Sydney, Australia

The reported experiment examines the impact of stem frequency on lexical decision responses to prefixed words. Both when the nonword distractors had nonsense stems (e.g., *recodge*) and real-word stems (e.g., *relaugh*), words with high frequency stems (e.g., *unreal*) were recognized more quickly than words with low frequency stems (e.g., *refuel*) when matched on surface frequency. This was taken as evidence that a whole-word representation exists for prefixed words, but that activation of this representation is always mediated by a representation of the stem, unlike the claims of a Dual Pathways model.

Keywords: orthographic processing, cross-language differences, syllabic processing, individual differences in reading, lexical processing

For the last decade or so, the most widely accepted view of affixed word recognition has been a Dual Pathways account whereby the word can be recognized either via a whole-word representation, or via a representation of its stem following decomposition (e.g., Baayen, Dijkstra, & Schreuder, 1997; Bertram, Laine, & Karvinen, 1999; Bertram, Schreuder, & Baayen, 2000; Colé, Beauvillain, & Segui, 1989; Niemi, Laine, & Tuominen, 1994; Niswander, Pollatsek, & Rayner, 2000; Schreuder & Baayen, 1995; Sereno & Jongman, 1997). This view contrasts with the Obligatory Decomposition account put forward by Taft for prefixed words (Taft, 1979; Taft & Forster, 1975; Taft, Hambly, & Kinoshita, 1986) and for inflected words (Taft, 1979, 2004), though it is really only the Taft (2004) study that directly tests Obligatory Decomposition against Dual Pathways. Because the Taft (2004) research was restricted to inflected words, the aim of the present experiment is to extend the examination of this issue to another type of affixed word, namely, prefixed words.

The Mental Lexicon 1:2 (2006), **183–199**.
ISSN 1871–1340 / E-ISSN 1871–1375

Evidence for the obligatory decomposition of inflected words (e.g., *jumped*, *mending*, *moons*), comes from a lexical decision task where the frequency of the stem of the word is manipulated (Taft, 2004). For example, while *seeming* and *mending* have the same frequency of occurrence in the language (i.e., the same "surface" frequency), the cumulative frequency (i.e., "base" frequency) of *seem*, *seemed*, *seems*, and *seeming* is higher than that of *mend*, *mended*, *mends*, and *mending*. Taft (2004) found that inflected words with a high base frequency (e.g., *seeming*, *moons*) were recognized more easily than those with a low base frequency (e.g., *mending*, *cliffs*), but only when those words were to be discriminated from nonwords that were inflected nonsense stems (e.g., *yaining*, *milphs*). When the distractor context in which the inflected words were placed consisted of inappropriately inflected real-word stems (e.g., *yearing*, *mirths*), not only did the advantage for the high base frequency words disappear, but such items were now so difficult to classify as words that a reverse base frequency effect emerged. That is, it became very hard to decide that *seeming* or *moons* was actually a word (as opposed to *mending* or *cliffs*) when the combination of the stem and affix had to be carefully considered.

The base frequency effect that was observed when the nonwords had nonsense stems (like *yaining*) implies that inflected words are decomposed into their morphemic constituents. If there were no decomposition, characteristics of the stem should have had no effect. The important question, though, is whether there also exists an alternative whole-word pathway that simply may not have been brought into play when the words and nonwords could be readily discriminated on the basis of their initial component (i.e., on the basis of the lexical status of their stem). The reverse base frequency effect that was observed when the nonwords had real-word stems (like *yearing*) strongly suggests that a whole-word pathway is not an option. If such a pathway were available, the introduction of such nonword distractors should have pushed the reader to rely on that pathway because information about the whole-word would then allow the words to be readily discriminated from the nonwords (i.e., *seeming* exists while *yearing* does not). This would result in the disappearance of the base frequency effect, but there would be no reason at all for high base frequency words to be harder to recognize than low base frequency words, as was observed.

On the other hand, a reverse base frequency effect can be readily handled by the Obligatory Decomposition account. The effect is ascribed to a late stage of processing whereby the functional information associated with the accessed representation of the stem morpheme is combined with the functional information associated with the accessed representation of the affix. The reason why a word with a high base frequency can be of relatively low frequency in the language (such as *seeming*) is often because the function of the affix does not readily accord with the function of the stem (see Taft, 2004). For example, *seem* is a stative verb that rarely takes the continuous aspect conveyed by the inflection *-ing*, unlike the dynamic verb *mend* which refers to something that occurs across time. As a result, the decision that *-ing* can combine with *seem* is harder to reach than the same decision for *mend*. What this means is that the advantage of accessing the high frequency stem *seem* over the lower frequency stem *mend* (i.e., the source of the base frequency effect) can be counterbalanced by the disadvantage when recombining *seem* and *-ing* relative to *mend* and *-ing*.

The degree to which such counterbalancing occurs will depend on the amount of weight placed on the recombination stage, and this will be influenced by the nature of the nonword distractors. When the nonwords have nonsense stems (e.g., *yaining*), a "yes" response can be potentially made on the basis of the accessed stem alone, with little need to consider the combinability of the stem and affix. On the other hand, when the nonwords have real-word stems (e.g., *yearing*), the only way to discriminate the words from the nonwords is on the basis of the combinability of the stem and affix. Therefore, the disadvantage at the recombination stage for low frequency words with high frequency stems (e.g., *seeming*) is accentuated under these conditions.

Now, the reason given for why the reader needs to consider the combinability of the function of the stem and affix of an inflected word is that no whole-word representation exists for such words. This is because all information associated with an inflected word can be predictably established purely on the basis of the information associated with its stem and its affix. For example, everything one wants to know about the inflected word *eating* can be entirely determined from the meaning of *eat* and the function of *-ing*. The same is not true for other types of affixed word, even though those words might be transparently derived from their stem and affix. For example, the meaning of the prefixed word *rebuild* is clearly derived from its verb stem *build* and the "again" meaning of its affix *re*. In contrast to the situation with inflected words, however, there is usually at least some information associated with a prefixed word that cannot be determined solely from information associated with its stem and affix alone. To illustrate, if one rebuilds a wall, it does not mean that one builds the wall in the same way that it was built in the first place, that is, from scratch. In order to capture such idiosyncratic information, there needs to be a representation of that whole prefixed word in lexical memory, a representation to which such information can be linked. Furthermore, such a whole-word representation must exist because we can distinguish real prefixed words from nonexistent prefixed words that nevertheless make sense (e.g., *regrill*, *unmerry*).

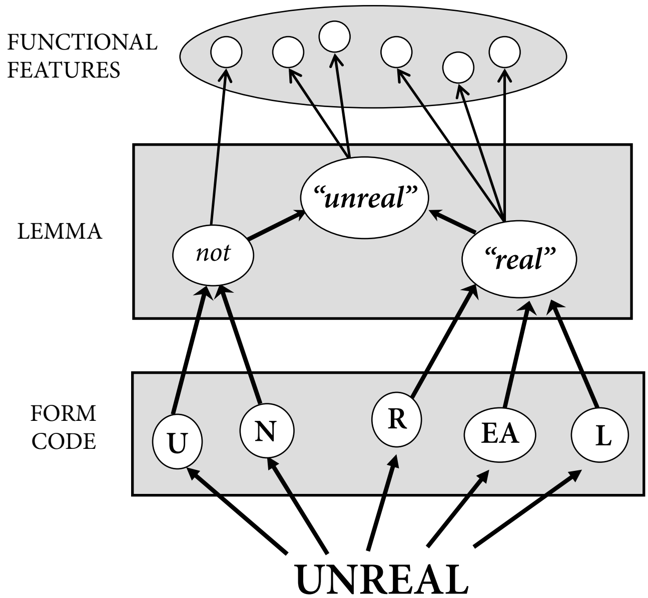

Figure 1. The representation of the prefixed word *unreal* in an Obligatory Decomposition model.

Just because there must be whole-word representations for prefixed words, however, it does not follow that a prefixed word must therefore be recognized without any decomposition into its affix and stem. It is still possible that the presented word is initially decomposed in order to activate separate representations for its prefix and stem, and it is only after this stage that the whole-word representation comes into play. In contrast to the recognition of inflected words, which requires a consideration of the way in which the function of the stem and affix combine together, the recognition of prefixed words can make use of the precompiled representation of that combination (i.e., the whole-word representation). The way in which there can be obligatory decomposition despite there being a whole-word representation is depicted in Figure 1.

According to this framework (see also Taft, 2003, 2004), information within the formal and functional systems is distributed across a number of different nodes. The form nodes are represented in Figure 1 by graphemes, but there could potentially be larger units than this, possibly hierarchically organized (see e.g., Taft, 1991; Taft & van Graan, 1998). Function nodes include both semantic and syntactic information, and are brought together with form nodes through the lemma level. A lemma is a unit that captures the correlation between form and function and, as such, can be seen as the lexical representation.

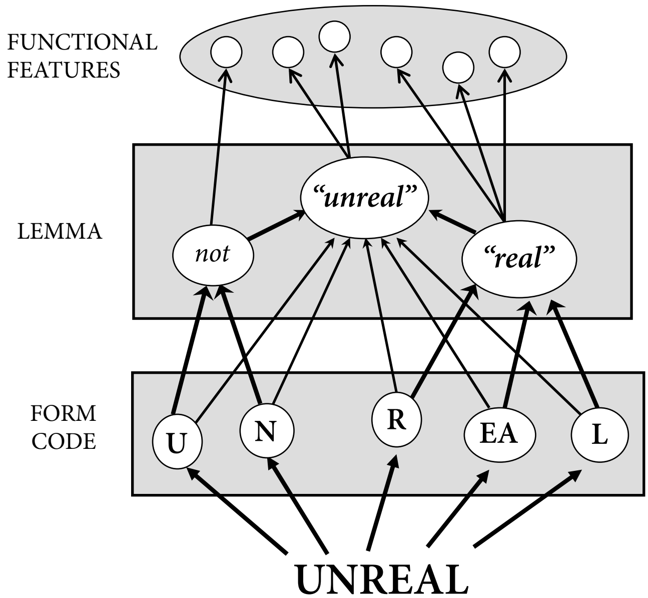

Figure 2. The representation of the prefixed word *unreal* in a Dual Pathways model.

It is proposed that there is a hierarchy of lemmas such that a prefixed word (e.g., *unreal*) is recognized via the activation of its stem (i.e., *real*) and its prefix (i.e., *un*). While much of the functional information relevant to the prefixed word might be activated through the lemmas for its stem and affix, there will be specific functional information that can only be captured by the lemma for the whole prefixed word (e.g., that *unreal* can refer to something amazing).

The important feature of Figure 1 for the present purposes, is the fact that there is a representation for the whole word, even though decomposition still occurs. In fact, decomposition is obligatory in this model because it is the only way in which the whole-word information can be reached. In contrast, Figure 2 illustrates a Dual Pathways account whereby the lemma for the prefixed word can be activated not only through decomposition, but directly from the form units as well. The latter pathway is what is usually referred to as "whole-word access", but within this framework, might be preferably termed "unmediated access" (given that a whole-word representation is also ultimately accessed via the decomposition procedure).

Is it possible to differentiate between these two accounts? In other words, is it possible to demonstrate that recognition of a prefixed word is always mediated by decomposition? As with inflected words, the impact of nonword distractors on the stem frequency effect can be used to address this issue. The stem frequency effect is where a word with a high frequency stem (e.g., *unreal*) is more readily classified as a word in a lexical decision task (on either reaction time [RT], error rates, or both) than one with a relatively low frequency stem (e.g., *refuel*)[1].

When the distractors in the lexical decision task are prefixed nonwords whose stem is nonsense (e.g., *recodge*, *undostle*), a real prefixed word can be differentiated from a nonword simply by establishing whether its stem exists in lexical memory or not. For example, *unreal* could be successfully classified as a word after the lemma for *real* is activated (see Figures 1 & 2). So if decomposition is used, the ease of activating the stem *real* will control the lexical decision response, thus generating a stem frequency effect. It must be noted, though, that the lemma for *unreal* could well be activated automatically once the lemma for *real* and *un* are activated and, therefore, the frequency of the surface form *unreal* could also have an impact. That is, while a stem frequency effect should emerge when surface frequency is controlled, it is quite possible that, in addition, a surface frequency effect will be observed when stem frequency is controlled.

The above prediction obviously holds for the Obligatory Decomposition account, but the Dual Pathways account predicts the same thing if the decomposition pathway dominates the unmediated pathway. If, on the other hand, the unmediated pathway is dominant, then only the whole-word representation would be accessed because there would be no need to proceed to the separate representations of the stem and prefix. Therefore, only surface frequency effects would be observed. So the presence of a stem frequency effect under these conditions (along with a concomitant surface frequency effect) would favor the importance of the decompositional pathway over the unmediated pathway in recognizing prefixed words.

The two theoretical accounts most clearly diverge when a nonword environment comprising prefixed real-word stems (e.g., *relaugh*, *unshout*) is considered. From the Obligatory Decomposition perspective, the prefixed word is still decomposed, but it is necessary to establish whether its accessed stem combines with its prefix, and this is based upon the existence of a lemma for the whole word. Thus, a stem frequency effect should still emerge when the distractors have real-word stems, though the surface frequency effect may increase in magnitude. The reason for this prediction is that activation of the whole-word lemma still passes through the stem representation (giving the stem frequency effect), but the whole-word lemma must always be activated (giving the surface frequency effect). When the nonword distractors have nonsense stems (like *recodge*) the whole-word lemma need not always be activated for a correct response to be made and, hence, the surface frequency effect may be weaker under those conditions. When the nonword distractors have real-word stems (like *relaugh*) mere activation of a lemma for the stem is insufficient for a correct decision to be made, so whole-word access must occur.

The Dual Pathway account makes a different prediction to the Obligatory Decomposition account when the nonwords have real-word stems (e.g., *relaugh*). If a stem frequency effect were observed when the nonwords did not have real-word stems (like *recodge*), it would suggest that the decomposition pathway typically dominates the unmediated pathway under those conditions. This might be taken to mean that the Dual Pathway account is virtually equivalent to the Obligatory Decomposition account, given that the unmediated pathway has so little importance. However, one could argue that decomposition dominates when the nonword distractors have nonsense stems simply because the words can be reliably discriminated from the nonwords on the basis of stem activation alone. If so, the introduction of nonword distractors that can only be discriminated from the words at the whole-word level (e.g., *relaugh*), should bring the unmediated pathway into greater prominence. That is, if direct whole-word access were available, it should become the dominant pathway when the task can best be performed on the basis of whole-word representations. Such an outcome would be evidenced by a loss of the stem frequency effect (as well as the strengthening of the surface frequency effect that is also predicted by the Obligatory Decomposition model).

So, to summarise, if a stem frequency effect were not found under either of the nonword environments, the Dual Pathways account (or even a nondecompositional account) would be supported. The Dual Pathways account could still be maintained if the stem frequency effect were found when the nonwords had nonsense stems (e.g., *recodge*), but it would lose its force if the effect remained when the nonwords had real word stems (e.g., *relaugh*). Finally, if the same pattern of results as that obtained with inflected words (Taft, 2004) were observed with prefixed words (i.e., a reversal of the stem frequency effect when the nonword distractors have real-word stems), it would have to be concluded that even prefixed words are recognized via a recombination of stem and affix based on functional information rather than on access to a whole-word representation.

The following experiment was designed to test the above predictions by examining the stem frequency effect under the two types of distractor conditions. A manipulation of surface frequency was also included. Since both accounts predict a larger surface frequency effect when the distractors have real-word stems, a failure to observe such an outcome would require a re-evaluation of the theoretical options.

Method

Participants. There were 34 participants who were all native English speakers studying first-year Psychology at the University of New South Wales. They were given course credit for their participation.

Materials. All of the experimental items were prefixed words with their stem being transparently related to the meaning of the whole word. Ten independent first-year Psychology students were asked to rate on a 7-point scale how much the meaning of the prefixed word was related to the meaning of its stem, and all words selected had ratings greater than 5.2. A range of different prefixes was used, though the majority were *un* or *re*. All frequencies were determined from the norms of Carroll, Davies, and Richman (1971), with the matching confirmed on the basis of Kučera and Francis (1967). Any words with frequency values that differed markedly between the two sets of norms were avoided.

To examine the stem frequency effect, 17 words with high frequency stems (HStem words, e.g., *unreal*, *rediscover*, *enforce*; mean stem frequency of 334 per million) were matched one-to-one with 17 words with lower frequency stems (LStem words, e.g., *refuel*, *relapse*, *unequal*; mean stem frequency of 30 per million[2]) on their surface frequency (0.67 and 0.63 per million respectively, $t < 1$). Length was matched over all pairs (mean length of 7.1 letters and 7.7 letters respectively, $t < 1.5$). Rated transparency (out of 7) was also matched, with ratings of 6.39 and 6.55 respectively ($t < .1$).

The surface frequency effect was tested with 17 words of high[3] surface frequency (HSurf words, e.g., *uncertain*, *rewrite*, *impossible*; mean surface frequency of 8.9 per million) and 17 words of low surface frequency (LS words, e.g., *repay*, *rework*, *uncommon*; mean surface frequency of 0.37 per million), matched one-to-one on their stem frequency (265 and 270 per million respectively[4], $t < 1$). Length was again matched over all pairs (mean length of 7.5 letters and 7.6 letters respectively, $t < 1$), as was rated transparency (6.76 and 6.62 respectively, $t < 1$).

In addition to the 68 experimental items, there were 45 nonwords that were also prefixed. Participants were randomly allocated to one of two equal-sized groups ($n = 17$), with each group receiving a different set of nonword distractors. For Group 1, the nonwords were all cases where the removal of the prefix produced a stem that was not a word in its own right (e.g., *recodge*, *untegure*, *debomp*, *regrondle*). These stems were all pronounceable strings of four to nine letters that were not similar to any real word that could genuinely combine with the prefix. In contrast, the nonwords presented to Group 2 were all cases where the removal of the prefix produced a real word (e.g., *relaugh*, *untragic*, *deburn*, *rebelong*). These real word stems extended across the full range of frequencies from very high frequency (e.g., *predifferent*) to very low (*refondle*), according to Carroll et al. (1971). The two sets of nonwords were comparable in terms of length and syllable structure.

All of the words and nonwords are found in the Appendix.

Procedure. Each participant was presented with a different random order of the 113 items, preceded by the same 10 practice items (where the nonwords corresponded to the type of nonword used in the actual experiment). Participants were told that they would see a series of letter-strings on the screen and were instructed to classify each of these as an existing word or nonexisting word as quickly but as accurately as possible by pressing a "yes" or "no" button. Each item was presented in lower-case letters under computer control. Stimuli remained on the screen until a response was made. If no response was made within 3 seconds, it was recorded as an error and the trial was terminated. There was an inter-trial interval of approximately 1 s.

Results

After any response over 2000 ms was eliminated, cutoff values for each participant were calculated for response times as two standard deviations above or below the mean across all correct "yes" responses. Any outlying values were replaced by the relevant cutoff value. Two participants in Group 2 (with real-stem distractors) were eliminated from the analysis owing to a mean error rate exceeding 25% for the word items. For the RT analysis, five item pairs were deleted owing to error rates of 25% or more for at least one member of the pair. Two of these pairs were in the stem frequency manipulation (*regrow/rekindle*, *subhuman/premature*) and three were in the surface frequency manipulation (*unaware/disunity*, *disagree/misdeal*, *rewrite/rework*). With such high error rates, RT becomes an unreliable measure, being based on too few responses. Mean lexical decision times and error rates are found in Table 1.

Because stem frequency (HStem vs. LStem) was manipulated independently of surface frequency (HSurf vs. LSurf), each was tested separately in a 2 × 2 ANOVA. The factors were relative frequency (higher vs. lower) and type of nonword distractor (non-stem vs. real-stem). In the analysis by participants (F_1), relative frequency was a within-group factor and nonword type was a between-groups factor. In the analysis by items (F_2), both factors were within-group, thus taking advantage of the pairwise matching of stimuli.

Table 1. Mean RT (in ms), % Error Rates (ER), and Item Standard Deviations (in parentheses) for the Stem and Surface Frequency Manipulations

		Non-stem distractors (e.g., *recodge*)		Real-stem distractors (e.g., *relaugh*)	
	Example	**RT**	**ER**	**RT**	**ER**
HStem	*rediscover*	549 (52.0)	7.62 (13.75)	667 (58.8)	14.12 (14.51)
LStem	*refuel*	580 (56.8)	8.71 (6.59)	704 (48.0)	15.31 (12.42)
Stem frequency effect		**+ 31**	**+ 1.09**	**+ 37**	**+ 1.19**
HSurf	*rewrite*	514 (33.1)	4.17 (3.47)	625 (28.4)	5.50 (4.85)
LSurf	*rework*	546 (49.6)	8.88 (14.02)	696 (76.1)	17.66 (19.85)
Surface frequency effect		**+ 32**	**+ 4.71**	**+ 71**	**+ 12.16**

Turning first to the stem frequency manipulation, the main effect of nonword distractor was significant for both RTs, $F_1(1, 30) = 5.63, p < .05; F_2(1, 14) = 135.38, p < .001$, and error rates, $F_1(1, 30) = 5.95, p < .05; F_2(1, 16) = 5.51, p < .05$, reflecting the greater difficulty in discriminating the words from the nonwords when the latter had real word stems. In terms of response times, there was a significant main effect of stem frequency, $F_1(1, 32) = 6.77, p < .02; F_2(1, 14) = 5.97, p < .05$, with no sign of an interaction with type of nonword environment, Fs < 1. Error rates showed no significant effects, all Fs < 1.

In relation to the surface frequency manipulation, again the main effect of distractor type was significant for RTs, $F_1(1, 30) = 9.05, p < .01; F_2(1, 13) = 144.30, p < .001$, and error rates, $F_1(1, 30) = 6.09, p < .02; F_2(1, 16) = 10.12, p < .01$. High surface frequency words were classified more quickly than low surface frequency words, $F_1(1, 30) = 23.33, p < .001; F_2(1, 13) = 10.44, p < .01$, and had lower error rates, $F_1(1, 30) = 22.04, p < .001; F_2(1, 16) = 4.26, p < .1$. The frequency effect was significantly larger in the real-stem distractor environment (e.g., *relaugh*) than in the non-stem distractor environment (e.g., *recodge*), $F_1(1, 30) = 4.42, p < .05; F_2(1, 16) = 5.45, p < .05$, but this interaction was only a trend for RTs, $F_1(1, 30) = 3.55, p < .1; F_2(1, 13) = 2.95, p > .1$. The surface frequency effect was significant for the real-stem group both on error rates, $F_1(1, 14) = 16.95, p < .001; F_2(1, 16) = 5.87, p < .05$, and on RTs, $F_1(1, 14) = 16.05, p < .001; F_2(1, 13) = 10.28, p < .01$. However, for the non-stem group, both measures were only significant in the participant analysis: $F_1(1, 16) = 4.76, p < .05; F_2(1, 16) = 1.67, p > .1$, for error rates, and $F_1(1, 16) = 6.44, p < .02; F_2(1, 13) = 3.35, p < .1$ for RTs.

Finally, responses to nonwords were far easier to make when the nonwords had a nonsense stem than a real-word stem both on RT (616 ms vs. 719 ms), $F_1(1, 32) = 18.31, p < .001; F_2(1, 88) = 51.89, p < .001$), and error rates (10.1% vs. 16.2%), $F_1(1, 32) = 7.50, p < .001; F_2(1, 88) = 6.87, p < .01$.

Discussion

The results support the Obligatory Decomposition account over the Dual Pathways account. A stem frequency effect is observed when the word items can be discriminated from the nonword distractors on the basis of their stem (e.g., *unreal* can be potentially discriminated from *recodge* on the grounds that there is a lemma for *real* but not for *codge*) and this suggests that the decompositional pathway is used under these circumstances. The Obligatory Decomposition explanation is that the lemma for the stem is activated because it is entailed by the processing of the prefixed word, whereas the Dual Pathways explanation is that the decompositional pathway dominates when the status of the stem is informative when performing the task. The fact that the stem frequency effect remains when the words can no longer be discriminated from the nonwords on the basis of the status of their stem (e.g., *unreal* vs. *relaugh*) is exactly what the Obligatory Decomposition account predicts. Unlike inflected words, there should be no disadvantage for the low frequency prefixed words with high frequency stems because there is no need to consider how well the stem recombines with the affix, given that there exists an actual representation that corresponds to the recombined form. For the Dual Pathways account, however, such a nonword environment is exactly where the unmediated pathway should take over from the decompositional pathway if the former pathway exists. The fact that the stem frequency effect remains for the RT measure suggests that the stem representation is still activated even under these circumstances, and it is hard to see why this would happen if an unmediated pathway were available.

One may wish to argue that decomposition is only dominant when activation of the whole-word lemma is very slow relative to activation of the stem lemma (as in the case of low frequency words with high frequency stems like *unreal*), but that the unmediated pathway would take over when it is the more efficient route to recognition as a result of relative frequencies. However, if true, the unmediated pathway would actually have very little use when it comes to prefixed words. Even for the highest frequency transparently prefixed words (e.g., *impossible*, *unhappy*), the stem is typically of much higher frequency than the whole word (e.g., *possible* is many times more frequent than *impossible*). In addition, few transparently prefixed words have a stem that is a very low frequency word, and when they do, the prefixed word tends to be even lower frequency than its stem (e.g., *rekindle*, *unlatch*). In other words, it is rare for the frequency of a prefixed word to be even close to that of its stem and, therefore, the decomposition pathway is going to provide the more rapid lemma activation in almost every case. So, even if an unmediated pathway were to exist for prefixed words, it would hardly ever come into play if a decompositional pathway also exists, and the RT findings of this study strongly support the idea that a decompositional pathway does exist.

The stem frequency data reported here with prefixed words contrast markedly with those obtained for inflected words, under conditions where the nonword distractors have real word stems (Taft, 2004). Inflected words show a strong reverse stem frequency effect under these conditions, whereas prefixed words maintain a recognition advantage for high frequency over low frequency stems. The explanation being given for the difference between inflected and prefixed words is that the former do not need a whole-word representation, while the latter do. That is, all information about an inflected word can be generated from stem and affix information alone, while this is not true for prefixed words, even when transparently derived from their stem and affix. It is apparent that prefixed words have a whole-word representation in lexical memory. However, the point being made in this study is that the existence of a whole-word representation does not imply that it is activated without any decomposition. Indeed, it is argued that such decomposition is obligatory or, at the very least, is the predominant method of recognition.

The findings in this study in relation to surface frequency are explicable by either the Obligatory Decomposition or the Dual Pathways account. There was a larger surface frequency effect in the real-stem environment (e.g., *relaugh*) than in the non-stem environment (e.g., *recodge*). The greater the need for whole-word information to discriminate the word items from the nonword distractors, the greater the impact of whole-word (i.e., surface) frequency, and this would be true for either account. In the case of the Obligatory Decomposition model, the fact that the distractors have a real-word stem (e.g., *relaugh*) forces processing to focus on the whole-word representation that has been activated via the individual representations of the stem and affix. On the other hand, the impact of such distractors for the Dual Pathways model is to bring the unmediated pathway into greater prominence relative to the decompositional pathway. In so doing, however, there should have been a concomitant reduction in the size of the stem frequency effect that arises from the decompositional pathway, but there was no hint of this happening. Thus, the Obligatory Decomposition account is favored.

Finally, mention should be made of an alternative model of morphological processing that has been presented in the literature, a Supralexical account (e.g., Giraudo & Grainger, 2000, 2003). According to this model, a presented polymorphemic word initially makes contact with a representation of the whole word, and this in turn activates its constituent morphemes, which feed activation back down to the whole-word representation. The whole-word representation is conceptualized in the Supralexical model as a form representation, implying that no decomposition need be carried out on the presented stimulus word. However, if it is accepted that the form representation comprises sublexical units, which are either graphemes (see Figure 1) or larger units (see e.g., Taft, 1991; Taft & van Graan, 1998), then the notion that the form of a presented word remains unanalyzed makes little sense. Instead, it can be suggested that the sublexical form units have no morphemic structure (as in Figure 1) and that the whole-word lemma is activated directly from these units, which leads to activation of the morpheme lemmas. Stem frequency effects would then be explained in terms of feedback to the whole-word representation from the lemma for its stem, and surface frequency effects would arise from the whole-word lemma itself. There are several problems with this approach, however.

First, it is unclear why the surface frequency effect should be greater when the nonwords change from having non-stems to real-stems. Feedback from the stem lemma is irrelevant to the surface frequency effect, so there is no source for any modulation of surface frequency arising from the nonword environment. Even more importantly, there is no reason for the nonword structure to be detected in the first place. In particular, the fact that a nonword (e.g., *relaugh*) has a real word as its stem will never be detected by a Supralexical system. That is, the prefixed nonword will fail to activate any whole-word representation, so there will be no pathway to its stem. If it were argued that the lemma for the real-word stem is partly activated by the sublexical form information, then the Supralexical model becomes indistinguishable from the Dual Pathways model because there will be concurrent activation of both the lemma for the whole word and the lemma for its stem.

It would also be hard to explain the results of Taft, Hambly, and Kinoshita (1986) which revealed a greater delay in classifying nonwords in lexical decision when the nonwords were composed of prefix plus bound stem (e.g., *dejoice* vs. *dejouse*) relative to when they were composed of a non prefix plus bound stem (e.g., *tejoice* vs. *tejouse*). This result strongly suggests that the morphemic status of the stem of a nonword is established through a decomposition process driven by the existence of a prefix, and such a conclusion is antithetical to the Supralexical model.

Conclusions

The study reported here supports the idea that prefixed words are recognized via a representation of their stem and affix. The frequency of the stem has an impact on lexical decision responses regardless of the type of distractor. That is, whether a prefixed word needs to be discriminated from nonwords with a nonsense stem or nonwords with a real-word stem, words with a high frequency stem are easier to recognize than words with a low frequency stem. Such a result is taken as evidence for obligatory decomposition in prefixed word recognition because it demonstrates the involvement of the stem even under circumstances where it would be advantageous to use a whole-word representation if one were to exist. The stem frequency effect obtained for prefixed words when the nonwords have a real-word stem stands in marked contrast to the reversal of this effect for inflected words (Taft, 2004). The conclusion is drawn that both types of word are decomposed, but recombination of the stem and prefix arises through activation of a whole-word representation (as in Figure 1), whereas recombination of the stem and inflectional suffix is based on functional information associated with each morpheme. It is expected that derivationally suffixed words should be treated in the same way as prefixed words unless, like inflected words, the meaning of the whole word can be entirely generated from the functions associated with their component morphemes (as may be the case with the derivational suffixes -*ness* and -*ly*, for example). Further exploration of the different types of polymorphemic word still awaits.

Author Note

The research reported in this paper was supported by a grant to the senior author from the Australian Research Council. The authors wish to thank Goran Lazendic for his assistance in carrying out the study.

Notes

1. In experiments with inflected words, the surface form is an inflectional variant of the stem and Taft (2004) used the term "base" frequency to refer to the cumulative frequency of all the inflectional variants of the stem to contrast it with surface frequency. On the other hand, when the surface form is a prefixed word, the cumulative frequency of all the inflectional variants of its stem no longer includes that surface form and, for this reason, the term "stem" frequency will be used here. Whether the surface frequency of a prefixed word should be included in calculating the cumulative frequency of its stem is immaterial when comparing prefixed words with high and low frequency stems because their surface frequencies are matched.

2. Although the stems in this condition are labeled "low frequency", they were mostly of medium frequency (as indicated by their mean frequency value of 30), with only two being genuinely rare (i.e., *canny* and *kindle*).

3. Since prefixed words are quite rare in English, "high frequency" refers here to a relative value.

4. Inclusion of the frequency of the prefixed word itself (along with its inflectional variants) in the calculation of stem frequency, made no impact on the matching of stem frequency between the HSurf and LSurf conditions. This is because surface frequencies of prefixed words are so low relative to their stem frequencies that they make only a very small contribution to stem frequency, even for the HSurf words.

References

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language, 37*, 94–117.

Bertram, R., Laine, M., & Karvinen, K. (1999). The interplay of word formation type, affixal homonymy, and productivity in lexical processing: Evidence from a morphological rich language. *Journal of Psycholinguistic Research*, *28*, 213–226.

Bertram, R., Schreuder, R, & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homophony, and productivity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26*, 489–511.

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

Colé, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language, 28*, 1–13.

Giraudo, H., & Grainger, J. (2000). Prime word frequency in masked morphological and orthographic priming. *Language and Cognitive Processes*, *15*, 421–444.

Giraudo, H., & Grainger, J. (2003). A supralexical model for French derivational morphology. In E. Assink, & D. Sandra (Eds.), *Reading complex words* (pp. 139–157). New York: Kluwer Academic/Plenum Publishers.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, R.I.: Brown University Press.

Niemi, J., Laine, M., & Tuominen, J. (1994). Cognitive morphology in Finnish: Foundations of a new model. *Language and Cognitive Processes, 9*, 423–446.

Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes*, *15*, 389–420.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale, NJ: Lawrence Erlbaum Assocs.

Sereno, J. A., & Jongman, A. (1997). Processing of English inflectional morphology. *Memory and Cognition, 25*, 425–437.

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

Taft, M. (1991). *Reading and the mental lexicon*. Hove, UK: Erlbaum.

Taft, M. (2003). Morphological representation as a correlation between form and meaning. In E. Assink, & D. Sandra (Eds.), *Reading complex words* (pp. 113–137). New York: Kluwer Academic/Plenum Publishers.

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 and Verbal Behavior, 14*, 638–647.

Taft, M., Hambly, G., & Kinoshita, S. (1986). Visual and auditory recognition of prefixed words. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 38A*, 351–366.

Taft, M., & van Graan, F. (1998). Lack of Phonological mediation in a semantic categorization task. *Journal of Memory and Language, 38*, 203–224.

Author's address:

Marcus Taft, School of Psychology, University of New South Wales, Sydney New South Wales 2052, AUSTRALIA. E-mail: m.taft@unsw.edu.au. Fax: 612–93853641

Appendix

The following are the prefixed words used in the experiment.

HStem/LStem pairs matched on surface frequency:

unreal/refuel, reappear/uncanny, rebirth/disobey, retell/reassemble, rediscover/recapture, reclaim/relapse, enforce/unequal, regrow/rekindle, refill/invalid, rethink/redial, unlock/immortal, unclean/disembark, recharge/misbehave, subhuman/premature, uncover/unsafe, retest/unromantic, unafraid/decode

HSurf/LSurf pairs matched on stem frequency:

rewrite/rework, informal/discredit, uneasy/redraw, rebuild/reopen, unaware/disunity, unload/reprint, incorrect/revisit, disagree/misdeal, uncertain/repay, disappear/subdivide, illegal/impolite, unkind/restart, unfair/unoriginal, uneven/unwell, incomplete/misunderstand, unhappy/undress, impossible/uncommon

The following are the non-stem distractors:

recodge, recammit, regrondle, unlaven, mislopple, disporfode, impezzate, debomp, untegure, rewoin, unslint, rebrommer, dischorgate, intalence, imprangle, enslemen, decrassle, precrofter, misramient, sublixent, illeminent, resquallen, untefferal, disrontle, intaffid, impestick, engroaf, derelore, predalfen, subjerdy, miscorrate, illexime, unaliagle, redellick, unscog, misterrian, disannate, incongish, remannote, unesta, impootle, undostle, reparrow, unwindle, misreppant

The following are the real-stem nonwords:

relaugh, rehappen, unshout, distolerate, engroan, unvalue, remoderate, disobtain, demuscle, resoft, refall, disprovide, enstamina, resquare, illisten, unsmoke, untragic, imbright, reterror, impaint, deburn, rejerky, imprivate, sublocal, incool, disjungle, insolid, unadmit, invacancy, misaccept, unearly, misanswer, unforeign, misramble, predifferent, debelong, implastic, prebangle, unhard, rebrook, redebate, unlemon, refondle, misdemand, unnature