LANGUAGE AND COGNITIVE PROCESSES, 1994, 9 (3) 271 294

# Interactive-activation as a Framework for Understanding Morphological Processing

Marcus Taft

*School of Psychology, University of New South Wales, Kensington, Australia*

A description is given of the main experiments that have been taken as support for the view that, in reading, a prefixed word is stripped of its prefix and lexically accessed on the basis of its stem. Since one of the most important of those experiments had been poorly executed, a new version of the same study is presented with results that are entirely consistent with the previous one. However, logical problems exist with the view that says that stems act as access codes used to gain access to the lexicon, the main ones having to do with the fact that a prefix store is required. As a result, an alternative model is favoured, namely, an interactive-activation model. Prefixed words are represented in decomposed form in this model, but no prelexical prefix-stripping is required. A detailed examination is made of the manner in which this framework is able to incorporate the previous empirical results, as well as other aspects of morphological processing.

## INTRODUCTION

What I wish to present in this paper is a discussion of the evolution of my thinking about morphological processing that has occurred over the last 20 years or so. My primary interest has centred upon words with bound stems (i.e. stems which are not words in their own right) which, in English, are typically prefixed words. The original framework in which I thought about the issue was that of a prelexical morphological decomposition procedure taking place prior to a search for the stem morpheme within the lexicon (Taft & Forster, 1975). In more recent times, however, I have favoured a

Requests for reprints should be addressed to Marcus Taft, School of Psychology, University of New South Wales, Kensington, New South Wales 2033, Australia. E-mail: M.Taft@unsw.edu.au.

The work reported in this paper was supported by a grant from the Australian Research Council. The author would like to acknowledge the technical help provided by Bruce Russell and Jishan Huang.

multilevel interactive-activation framework, where one of the levels contains morphemic units (Taft, 1991). I am by no means the first to propose a model along these lines (see, e.g. Andrews, 1986; Dell, 1986; Fowler, Napps, & Feldman, 1985; Grainger, Colé, & Segui, 1991), but what I hope to provide here is a more detailed examination of what such an interactive-activation account requires in order to explain the previous experimental results that have been taken as support for the original model. First, though, I will give a brief historical account of my original position.

## ACCESSING PREFIXED WORDS VIA THEIR BOUND STEMS

A single item in an experiment I was running in 1974 led me to the idea that the mental lexicon might in some way represent the bound stems of prefixed words. In that experiment, subjects were presented with nonwords that were homophonic with real words, and were asked to give the meaning of the homophonic word. What I noticed was that a very frequent response to the item SKRYBE was to say that it had "something to do with writing". I found it rather surprising that so many subjects referred to the "writing" aspect of SCRIBE without mentioning the important fact that it refers to a person (whose job it is to write). This got me thinking that their interpretation of SCRIBE was actually drawing upon the various words which include that word as a bound stem, like INSCRIBE, DESCRIBE, ASCRIBE, SUBSCRIBE, as well as SCRIBE itself, and possibly even SCRIBBLE and SCRIPT (cf. INSCRIPTION, DESCRIPTION, etc.), where the unifying theme might be taken to be "writing". It was this possibility that led to the series of experiments that was subsequently published in the 1975 Taft and Forster paper.

The central experiment in that paper was one where subjects were presented with letter strings and were required to decide whether each was a word or not (i.e. lexical decision). Of most interest were the responses to nonwords which were either the bound stems of prefixed words (i.e. VIVE from REVIVE), the "pseudo-stems" of pseudoprefixed words (e.g. LISH from RELISH), and the last parts of non-prefixed words (e.g. NACE from MENACE). On finding that real stems took longer to classify as nonwords than either pseudo-stems or non-stems (which did not differ from each other), it was concluded that bound stems are indeed given some sort of lexical representation. It is an encounter with that lexical representation that leads to a delay in making a nonword response to a real stem item.

Now, I have often encountered criticism of this experiment, usually by word of mouth rather than directly in print. Henderson (1985; 1986) reviews the issue of morphological processing, but pointedly eschews discussion of this experiment. There seem to be two general criticisms:



first, that the materials were poorly controlled, and second, that the stimuli were nonwords rather than words. I will now attempt to address these concerns, since the conclusion that bound stems are lexically represented is an important one.

I acknowledge the fact that the materials used in the Taft and Forster (1975) experiment were very poorly controlled. Those were early days for research into lexical processing, and that experiment may indeed have been the first to use a design whereby items were matched across conditions on a one-to-one basis taking into account factors such as length and frequency (e.g. VIVE, LISH and NACE all have four letters, and the words from which they are derived have the same word frequency). However, we now know that controlling for length and base-word frequency are not sufficient to control for the word-likeness of nonwords. For example, Coltheart, Davelaar, Jonasson and Besner (1977) demonstrated that the more words that are one letter different from a nonword (i.e. the larger the "neighbourhood"), the longer the lexical decision responses to that nonword. In addition, most of the real stems used in the experiment coincided with the stressed part of the base word (e.g. the VIVE of REVIVE), whereas the pseudo- and non-stems did not (e.g. the NACE of MENACE), and also the number of words containing the nonword segment was not controlled. Finally, the decision about whether a stem was real or pseudo was based primarily on intuition, which apart from anything else could have led to a bias in classifying an item as a real stem because of an anticipated difficulty in classifying it as a nonword.

## A NEW EXPERIMENT

Recently, I repeated the experiment with an improvement in the construction of the materials. Twenty-two real stems (e.g. PECCABLE from IMPECCABLE, TENUATE from ATTENUATE, THRALL from ENTHRALL, NIHILATE from ANNIHILATE, SCURE from OBSCURE, COCIOUS from PRECOCIOUS) were compared with 22 nonwords which did not form the ending of any real word, but which were matched to a real stem on as many factors as could be thought of in order to control for word-likeness (e.g. TUCCABLE, ARDUATE, SHRALL, SOLECATE, STURE, TOTIOUS). The stems and control nonwords were matched on the number of words which could be formed by changing one letter. Other examples of how word-likeness was tackled were things like the PECC of PECCABLE being pronounced like PECK and the TUCC of TUCCABLE being pronounced like the equally frequent word TUCK, both TENUATE and ARDUATE forming equally frequent words when the -ATE is replaced by -OUS, and both NIHILATE and SOLECATE forming equally frequent words when -ATE is replaced by -ISM.

Half of the real stems had a higher bigram frequency than their control nonwords and half had a lower bigram frequency (from Mayzner & Tresselt, 1965). Classification as a real stem was based on the ratings of 10 independent judges who were asked to rate the prefixedness of a large corpus of words. Words with an average rating above 4 out of 7 were classified as prefixed, with the average rating for the 22 selected items being 4.7.

In addition to these two sets of nonwords, 22 non-stems were selected (e.g. TINERARY from ITINERARY, ZETTE from GAZETTE, BORIOUS from LABORIOUS, NIPULATE from MANIPULATE, BITUATE from HABITUATE, REER from CAREER) and these were similarly matched on word-likeness to 22 other nonwords which did not form the endings of any real word (e.g. GINERARY, ZOTTE, MORIOUS, DABULATE, CATUATE, GEER). Each non-stem was paired with the real stem on approximate length and general structure (e.g. PECCABLE/TINERARY, TENUATE/BITUATE, NIHILATE/NIPULATE, COCIOUS/BORIOUS). The set of non-stems was matched overall with the set of real stems in a number of ways. Both always formed the stressed segment of a whole word, and that was the only word in which they occurred. The frequency of these words was approximately matched across the real stem and non-stem items, as was the type-frequency of the letters that were deleted to form the nonword (according to the Word-perfect 5.1 computer word-processing software lexicon). As with the real stems, the non-stems and their controls were matched overall on bigram frequency.

Because both the real stems and the non-stems were very similar in structure to their matched controls, two groups of subjects were used, each receiving one member of any stem/control pair (e.g. either PECCABLE or TUCCABLE, TINERARY or GINERARY). Thus each subject saw 11 items in each of the four conditions. There were 15 native English speakers in each group. The 44 nonwords seen by each subject were interspersed with 44 real words of similar structure (e.g. DINOSAUR, FREAK, CAPABLE). The real words and the words from which the nonwords were derived came from a similar frequency range.

Even with the rigorous construction of items, the results of this lexical decision experiment upheld the conclusions drawn from the original Taft and Forster study. The results are presented in Table 1. The interaction between real stem versus control and non-stem versus control was significant, both on reaction time [$F_s(1,28) = 10.426$, $P < 0.01$; $F_i(1,21) = 13.868$, $P < 0.01$] and errors [$F_s(1,28) = 7.891$, $P < 0.01$; $F_i(1,21) = 4.786$, $P < 0.05$]. This interaction reflected the fact that real stems significantly differed from control nonwords on both reaction time [$F_s(1,28) = 14.58$, $P < 0.001$; $F_i(1,21) = 18.58$, $P < 0.001$] and on errors [$F_s(1,28) = 23.87$, $P < 0.001$; $F_i(1,21) = 5.75$, $P < 0.05$], while the non-stems did not differ at all from their controls on either reaction time or error rates (all $F$'s < 1).

So we see that the conclusions drawn by Taft and Forster (1975) are confirmed with more carefully controlled stimuli. Nonwords which are considered to be stem morphemes are treated as being more word-like than those which are not, and this suggests that such stem morphemes in some way provide automatic access to the words of which they are a part. According to Taft and Forster, this comes about because prefixed words are actually represented by their stem in the lexical access system. The stem is seen as being an "access code", through which one gains access to information about the full word. In order to discover this access code within a presented prefixed word, the prefix must be stripped off so that the stem can be isolated.



TABLE 1
Lexical Decision Times (msec) and Error Rates for Nonwords Used in the Improved Version of the Experiment Performed by Taft and Forster (1975)

| *Condition* | *Example* | *RT* | *Errors (%)* |
|---|---|---|---|
| Real stems | PECCABLE | 1262 | 20.0 |
| Control | TUCCABLE | 1120 | 9.1 |
| | difference | 142 | 10.9 |
| Non-stems | TINERARY | 1131 | 10.0 |
| Control | GINERARY | 1151 | 7.9 |
| | difference | −20 | 2.1 |

Evidence for such a prefix-stripping procedure comes from a study by Taft, Hambly and Kinoshita (1986), which demonstrated a delay in nonword decisions when the nonword included a real stem, but only when the stem was preceded by a prefix. For example, INVIVE took much longer to respond to than INVAPE, whereas IBVIVE did not differ from IBVAPE. That is, the real stem VIVE was able to delay nonword decisions only when it could be isolated via prefix stripping. IB is not a prefix, and was therefore not stripped, so that the existence of the real stem VIVE after the IB had no impact on response times.

## USING NONWORDS AS STIMULI

As mentioned earlier, experiments of the type mentioned so far have been criticised on the grounds that they examine responses to nonwords, which might involve special strategies that are irrelevant to normal reading (e.g.

Henderson, 1986). While it may be true that classifying a letter string as a nonword requires extra processing than classifying one as a word, it would be unjustifiable to dismiss the difference observed between real stems and non-stems as being the result of an artificial strategy. The fact that PECCABLE is in some sense more word-like than TINERARY would seem to have something to do with lexical representation rather than some strategically manipulable retrieval mechanism. What makes PECCABLE confusable with IMPECCABLE while TINERARY is not confusable with ITINERARY can only be the fact that PECCABLE coincides with the morphemic structure of IMPECCABLE, whereas TINERARY does not coincide wtih the morphemic structure of ITINERARY.

It might be argued, though, that the realisation of this coincidence with morphemic structure arises only *after* the word has been accessed. That is, PECCABLE and TINERARY might both be able to access the respective words of which they are a part on the basis of their orthographic similarity to those words, but information about morphemic structure contained within the accessed lexical entries suggests that PECCABLE is a potentially acceptable structure, whereas TINERARY is not. Such a view would hold that morphemic structure plays no role in lexical access itself. However, if this were so, it would seem that there should be some difficulty evidenced in responding to TINERARY compared with GINERARY, because the former gains access to a lexical representation, whereas the latter does not. Since there is no such difficulty, it seems that while PECCABLE gains access to information about IMPECCABLE, TINERARY does not gain access to information about ITINERARY (at least within the time framework of the lexical decision task). It is for this reason that the difficulties observed with real stem nonwords would seem to be reflecting the manner in which prefixed words are structured for the purposes of retrieval.

## EXPERIMENTS USING WORDS AS STIMULI

Notwithstanding these arguments in favour of the relevance to normal lexical processing of experiments which examine responses to nonwords, there are other experiments which have been taken as support for the Taft and Forster model that use words as stimuli. The content of these experiments has been documented elsewhere (e.g. Taft, 1985; 1988; 1991), so I will only outline them briefly here:

1. If a word can be the bound stem of a higher frequency word, it takes longer to classify as a word than one which cannot (Taft & Forster, 1975). For example, VENT is associated with longer lexical decision times than the frequency-matched word COIN, where VENT is the bound stem of the higher frequency words INVENT and PREVENT.[1] According to the Taft and Forster model, this result comes about because VENT has two representations in the access system, one for the purpose of accessing INVENT and PREVENT, and the other for the purpose of accessing the word VENT itself. Because the former is more common than the latter, it is accessed first and thus interferes with access to the appropriate entry that would allow classification as a word.

2. Lexical decision times are influenced by the frequency of the bound stem of a prefixed word when whole word frequency is controlled (Taft, 1979a). For example, while REPROACH and DISSUADE have the same frequency of occurrence, the stem PROACH is more common than the stem SUADE because of the frequency of the other word which contains that stem. That is, APPROACH is more common than PERSUADE. Words like REPROACH are associated with faster lexical decision times than words like DISSUADE. Such a result is consistent with the idea that access to lexical information about a prefixed word takes place through a representation of its stem.

3. Non-prefixed words which begin with letters that can function in other words as a prefix (i.e. "pseudoprefixed" words like REGATTA, DISCIPLE or PREDATION) take longer to process than non-prefixed words like GRAFFITI, CADET or TABASCO, which do not begin with such letters (Bergman, Hudson, & Eling, 1988; Lima, 1987; Taft, 1981). The Taft and Forster model predicts such an outcome, since pseudo prefixes should be inadvertently stripped off prior to access, and this will turn out to be an inappropriate retrieval strategy for non-prefixed words. REGATTA cannot be accessed via a representation for GATTA, since it is not a stem morpheme, and any attempt to do so will slow down access times.

## PROBLEMS FOR THE TAFT AND FORSTER MODEL

The idea that a prefixed word is stripped of its prefixes and the stem used as an access code for the whole word, allows for little flexibility in processing. One might suppose, therefore, that such a strong theory would be readily falsifiable. While I believe that there are actually no results which have clearly falsified it, there are certainly results which reduce its credibility.

If the finding of Andrews (1986) were confirmed that stem frequency effects can come and go depending on the experimental context, this would be a major setback to the Taft and Forster position. Stem-frequency effects are supposed to arise from the nature of the stem representation, not from

[1]This experiment is a formalisation of the anecdotally reported difficulties that I observed with the word SCRIBE.

some modifiable retrieval mechanism. Andrews observed a stem-frequency effect with derived words only when compound words were included in the experiment. Against this, however, is the fact that Bradley (1979) was able to obtain stem-frequency effects with derived words when no compound words were included in the study, and no additional items which might have induced the stem-frequency effect were used in the Taft (1979a) experiment with prefixed words.

Recent demonstrations that semantic factors influence morphological processing (e.g. Marslen-Wilson, Tyler, Waksler, & Older, 1994; Sandra, 1990; Zwitserlood, this issue; but see Bentin & Feldman, 1990) would seem to go against the spirit of the Taft and Forster model, which focuses on the nature of a semantic-free access system. However, it is possible to say that whether or not a stem is used as the access code for a whole word is influenced by semantic factors. Indeed, such a position was required in differentiating the storage of prefixed words and pseudoprefixed words. This in itself, though, is unsatisfying, since there is no particular reason for semantic factors to play a role within a code which exists purely as a means of gaining access to a whole word. For example, there is actually no particular reason why REGATTA should not be accessible through the pseudo-stem GATTA, since the prefix is going to be stripped off anyway.

The requirements that there be a "prefix store" is another aspect of the model which detracts from its appeal. If prefixes are to be stripped off prior to access, there must exist a prefix store which allows recognition of putative prefixes independent of other lexical information. There are at least two problems with this. First, it is not obvious that readers know what are potential prefixes and what are not. In fact, many years ago I attempted to run an experiment in which subjects were asked to decide whether a letter string was a prefix or not (e.g. "yes" to IM, EX, DIS; "no" to IB, AX, FAM), but found that most subjects were quite unable to perform the task. If there exists an independent list of prefixes, one might have expected it to be quite easy to differentiate a prefix from a non-prefix. Second, there are some letter strings which have a dubious status as far as being a prefix goes. These are initial "combining forms" like POLY, BIO, MINI, PSEUDO and JURIS. Should such forms be considered to be a prefix and therefore be stripped (thus capturing the relationship between JURISDICTION and PREDICTION, for example), or should they be treated like one of the stems of a compound word and thus be an access code in their own right? If the former, the prefix store would have to be quite amorphous: For example, if BIO were considered to be a prefix, then so should PSYCHO, ANTHROPO, CRIMINO, METHODO, etc. Although this seems far-fetched, I have found that many people will rate words like JURISDICTION, VERDICT and VIADUCT as being prefixed words, and this seems incompatible with the idea of a clearly delimited store of prefixes.



## AN INTERACTIVE-ACTIVATION MODEL

It is partly for reasons along these lines that I have begun to favour an interactive-activation model (see, e.g. McClelland & Rumelhart, 1981; McClelland, 1987) as a framework for thinking about morphological processing. In such a model, there is no prelexical prefix stripping, so there need not be any specific store of prefixes. Prefixes are nevertheless treated separately from their stems, by virtue of the fact that they constitute independent activation units.

The other reason for postulating such a framework is that it is a useful way of handling other aspects of lexical processing, particularly the involvement of phonology in accessing visually presented words (see Taft, 1991). Not only is lexical processing influenced by orthographic–phonological relationships at the single grapheme/phoneme level, but also by those at the level of the word body (i.e. that part of a monosyllabic word that does not include the initial consonants), like the INT of PINT or the UTE of FLUTE. The phonological equivalent of a body is a "rhyme", and one body can be associated with more than one rhyme (e.g. the INT of PINT and MINT) and vice versa (e.g. the /u:t/ of FLUTE and SHOOT). One can capture this fact by postulating units of activation, both orthographic and phonological, at the single segment level as well as at the body level (see Taft, 1991, p. 81). A simplified schematic version of this is given in Fig. 1.

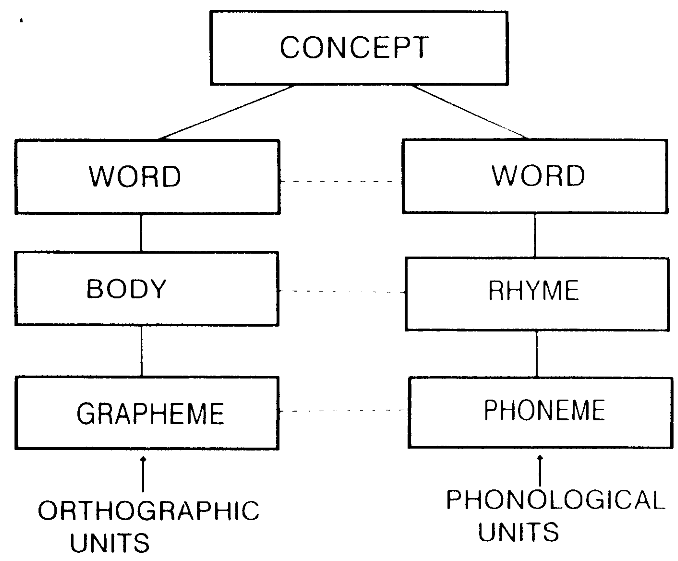


FIG. 1 A version of the interactive-activation model.

The "concept" units constitute the interface between the representation of form and meaning. When activated through the lower level units, these units make available the semantic characteristics of the word. The orthographic and phonological units representing the whole word are excluded from the version of the model presented in Taft (1991), since they were considered unnecessary. However, they are included here for the sake of clarity.

In considering how a morpheme level fits into this interactive-activation framework, we will focus solely on the orthographic units for the moment. What I want to suggest (see also Taft & Zhu, 1994) is that there are morpheme units which intervene between the body level and whole word level, which may or may not themselves be associated directly with concept units. Such an idea is presented in Fig. 2, where the prefixed word INVENT is used as an example.

Activation enters the system via the grapheme units and works its way up to the concept level, though during this process increased activation at higher levels can feed back down to enhance the activation of promising lower level units. The frequency with which a pathway to a unit is used will influence the strength of activation within that unit, such that a unit representing a high-frequency word will be activated more strongly than one representing a low-frequency word. A lexical decision response can be

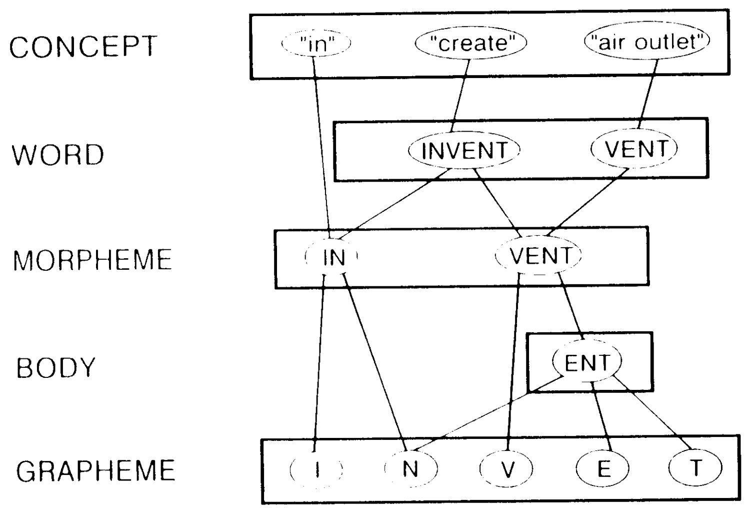


FIG. 2 The interactive-activation model incorporating a morpheme level, depicting the representation for INVENT.



made on the basis of the amount of activation at the whole word level,[2] though the time taken to reach a decision will depend on the amount of competitive activation in the other units.

In the case of INVENT, we see that activation of the stem VENT comes about through activation of the grapheme nodes V, E, N and T, via the body node ENT. Activation from this morpheme node for VENT in turn activates the word units representing both INVENT and VENT (as well as PREVENT, ADVENT, CONVENT, VENTURE, etc.). At the same time, the morpheme node IN becomes activated through the grapheme units I and N, and as such helps to raise the activation level of the word node INVENT beyond that of the competing candidates. Note that the morpheme unit IN is directly linked to a concept node, since many words which begin with IN seem to share a sense of "in-ness" (even though this is not obvious in the word INVENT). No direct conceptual connection is included in Fig. 2 for the morpheme VENT, since it is less obvious what concept is shared by those words which include it. However, should the reader be able to abstract out such a shared sense (maybe being familiar with the Latin origin, "come"), then their lexicon may well have a direct connection between the morpheme VENT and the concept level.[3]

## EXPLAINING THE DATA

Let us now examine how a framework of this type can account for the results of previous research which were taken to be support for the prefix-stripping/search model of Taft and Forster.

*Bound Stem Nonwords.* It is easy to explain the finding that bound stems like VIVE (or PECCABLE) take longer to classify as nonwords than pseudo-stems like LISH and non-stems like NACE (or TINERARY). At the morpheme level, REVIVE will be represented as RE and VIVE, while RELISH and MENACE, being monomorphemic, will only be represented at the word level (or perhaps at both the word level and morpheme level). Therefore, when VIVE is presented, there will be activation of the VIVE node, which in turn will partially activate the word node REVIVE. If there is sufficient activation in that word node, an erroneous "yes" response will

[2]If the whole word level were eliminated from the model (as suggested by Taft, 1991), the lexical decision response would be based on the amount of activation at the concept level. The morpheme units IN and VENT, for example, would converge on the "create" concept unit rather than the "INVENT" word unit.

[3]Note that other variations on Fig. 2 are possible. For example, instead of activation passing to the word unit VENT from the morpheme unit VENT, it might pass directly from the body unit ENT and grapheme unit V.

be made but if such an error can be avoided (see Taft, 1991, for a discussion of how this might occur), then the time taken to respond correctly will be delayed. On the other hand, neither LISH nor NACE will activate any particular sublexical node and therefore activation at the word level will be minimal.

*Bound Stem Words.* An explanation along similar lines can be given for the delay observed in responding to bound stems which happen to be words as well (like VENT compared with COIN). As can be seen from Fig. 2, when VENT is presented, activation in the VENT word node will be in competition with activation in the INVENT word node and, therefore, the response will take longer than that to a word for which there is no such competition (like COIN).

*Stem Frequency.* The stem-frequency is explained in the same way that the word-frequency effect is explained, except in terms of morpheme units rather than word units. That is, a high-frequency morpheme node (e.g. PROACH) is activated more strongly than a lower frequency morpheme node (e.g. SUADE), though the strength of association between the morpheme node PROACH and the word node REPROACH is the same as that between SUADE and DISSUADE. The stem-frequency effect therefore arises from the advantage in activation at the morpheme level for the word with the higher frequency stem.

*Pseudoprefixed Words.* Pseudoprefixed words like REGATTA can take longer to process than non-prefixed words like GRAFFITI because of the competition arising from the activation of an inappropriate node. In particular, the morpheme node RE will be activated (along with the concept node "again"), which in turn will lead to competition with the REGATTA node at the word level and at the concept level (as well as at the morpheme level if a monomorphemic word is represented at both the morpheme and word levels).

*Non-prefixed Stem Nonwords.* The finding that bound stems do not delay lexical decision responses when preceded by a non-prefix (e.g. IBVIVE = IBVAPE) is not so easy to explain. In an activation model, nodes will become activated automatically in the presence of the appropriate stimuli. Therefore, VIVE should arouse activation in the VIVE morpheme unit no matter what precedes it. The fact that Taft et al. (1986) found no effect here seems more in line with the notion that stems are only isolated via an active prefix-stripping procedure. One possible way to explain the result in terms of interactive-activation is to introduce a strong left-to-right bias into the system. In particular, if the first few letters of the



stimulus fail to activate any morpheme or word units to some criterial level, then the system can assume that the stimulus is a nonword and respond accordingly at that point. In this way, the remaining letters will only have an impact if they can activate units sufficiently at the morpheme or word level before the decision is made to opt out. For this reason, it might be possible to find nonword items which do produce a delay in lexical decision times even though they do not begin with a prefix—for example, when the final letters form a common morpheme or word which is rapidly activated (e.g. IBPEOPLE *vs* IBPEADLE).

## FURTHER ISSUES TO CONSIDER

### A New Definition of Prefix Stripping

The interactive-activation account is more flexible than the prefix-stripping/search framework, in that the equivalent of the prefix-stripping procedure is an integral part of the access process itself rather than a discrete stage of processing that takes place prior to acces. A prefix always participates in the accessing of the whole word representation, but its impact on the outcome of the accessing process is variable. If it is a very common prefix, it will serve to activate so many candidate words that it will turn out to be very uninformative. In this case, the stem will play the major role in accessing the word representation, and this is functionally equivalent to prefix stripping. When the prefix is relatively uncommon, however, that prefix will play more of a role in accessing the word. In this way, specific knowledge about whether or not a letter sequence actually forms a prefix is unimportant for the purpose of accessing a word.

In fact, such knowledge is not even required for the word to be stored with a prefix/stem structure in the first place. Such a structure would simply emerge from the relationships that exist between words which share the same letter sequences. For example, INVENT might be represented at the morpheme level as IN and VENT, because IN is a potential prefix owing to its recurrence in many words which share the concept of "in-ness" (e.g. INGRAIN, INSTIL, INDULGE, INGRATIATE) and the remaining letter sequence VENT recurs in a number of words which also begin with putative prefixes (e.g. PREVENT, ADVENT, CONVENT, CIRCUMVENT).[4] The amount of overlap between the meaning associ-

[4]There would need to be somewhat more sophisticated knowledge required in addition to these patterns of letter sequences. For example, one would need to know whether a bound morpheme functions as a stem or a prefix (perhaps through position-sensitive analysis). For example, even though FER might exist as a bound morpheme (cf. REFER, DEFER, TRANSFER, INFER, etc.), the word FERVENT should not be analysed as FER plus VENT because FER does not function as a prefix.

ated with a prefix and that associated with the word in which it occurs is also likely to play a role in determining morphological structure. For example, even thought PECCABLE occurs in no words other than IMPECCABLE, the "not" meaning of IM appears to be sufficiently contained in the "no mistakes" meaning of IMPECCABLE for that word to be considered as prefixed.

## Prefixed Words with Free Stems

How does the model incorporate words, like REHEAT, whose stem is a word which is transparently related to the prefixed form? Since HEAT is both a word and a morpheme, one might suppose that it would be represented at both those levels, while REHEAT would be represented at the word level only. Such a view is depicted in Fig. 3.

There are several problems with such a conceptualisation. First, there appears to be unnecesary redundancy not only in the repeated representation of HEAT, but also in the fact that the concept node associated with REHEAT is merely a concatenation of the separate concept nodes associated with RE and HEAT, and this relationship is not captured by the model. Second, since the concept node for REHEAT is quite independent

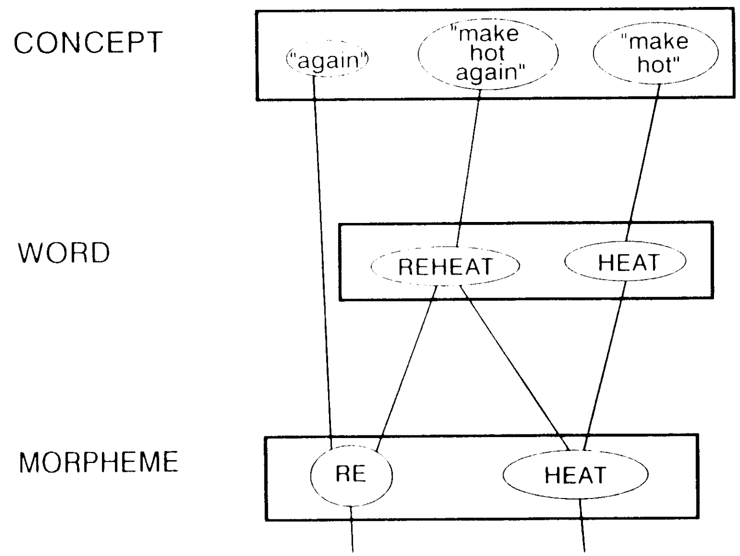


FIG. 3 A possible way of depicting the representation of prefixed words using REHEAT as an example.



of the concept nodes for RE and HEAT, such a set-up does not provide any distinction between semantically transparent and semantically opaque forms (cf. Marslen-Wilson et al., 1994; Sandra, 1990; Zwitserlood, this issue). For example, RELATE would be associated with a concept node "pertain to", which has nothing to do with the concept node associated with LATE. Yet, the nodal structure would be exactly the same for transparent words (like REHEAT) and opaque words (like RELATE).[5]

Is there a way, then, to avoid redundancy while also capturing the difference between transparent and opaque forms? Figures 4a and b illustrate one possibility. In Fig. 4a we see how the transparent word REHEAT might be represented, and Fig. 4b depicts the same for the opaque word RELATE.[6] In this version of the model, the only units found at the morpheme level are bound morphemes. Within the word level, monomorphemic words (e.g. HEAT) can be connected to polymorphemic words (e.g. REHEAT), and this is really equivalent to a division of the word level into separate monomorpheme and polymorpheme sub-levels. Such a version of the interactive-activation model has also been advocated by Taft and Zhu (1994) on quite different grounds.

The difference between transparent and opaque words is two-fold. First, the word which coincides with the stem of an opaque prefixed word (e.g. the word LATE coinciding with the stem of RELATE), is not connected to the node representing the prefixed word. Second, the word node representing a transparent prefixed word is not associated with a concept node, whereas that representing an opaque prefixed word is. In this way, the combined activation of "again" and "make hot" is used to construct the meaning of REHEAT, while activation in the "pertain to" concept node will dominate any competing activation of "again" and "not on time" when RELATE is presented.

The reason for including a word node for the transparent prefixed word at all, is to differentiate it from a prefixed neologism which the reader has never encountered before, but can nevertheless understand. For example, it is easy to work out what REGRILL means, even though, unlike REHEAT, it is not a real word for most people. In terms of Fig. 4a, the meaning can be determined from the combined activation of "again" and "cook on a griller", but there is no actual word unit representing

[5]A distinction must be made between transparent and opaque forms if for no other reason than the fact that RE is generally pronounced differently when used in a transparent form than when used in an opaque form (/riː/ for transparent, cf. REHEAT; /rə/ for opaque, cf. RELATE).

[6]Note that the only changes required of Fig. 2 to make it compatible with Fig. 4 are first, that the morpheme level becomes specifically a "bound morpheme" level, and second, that the variation suggested in Footnote 3 is implemented.

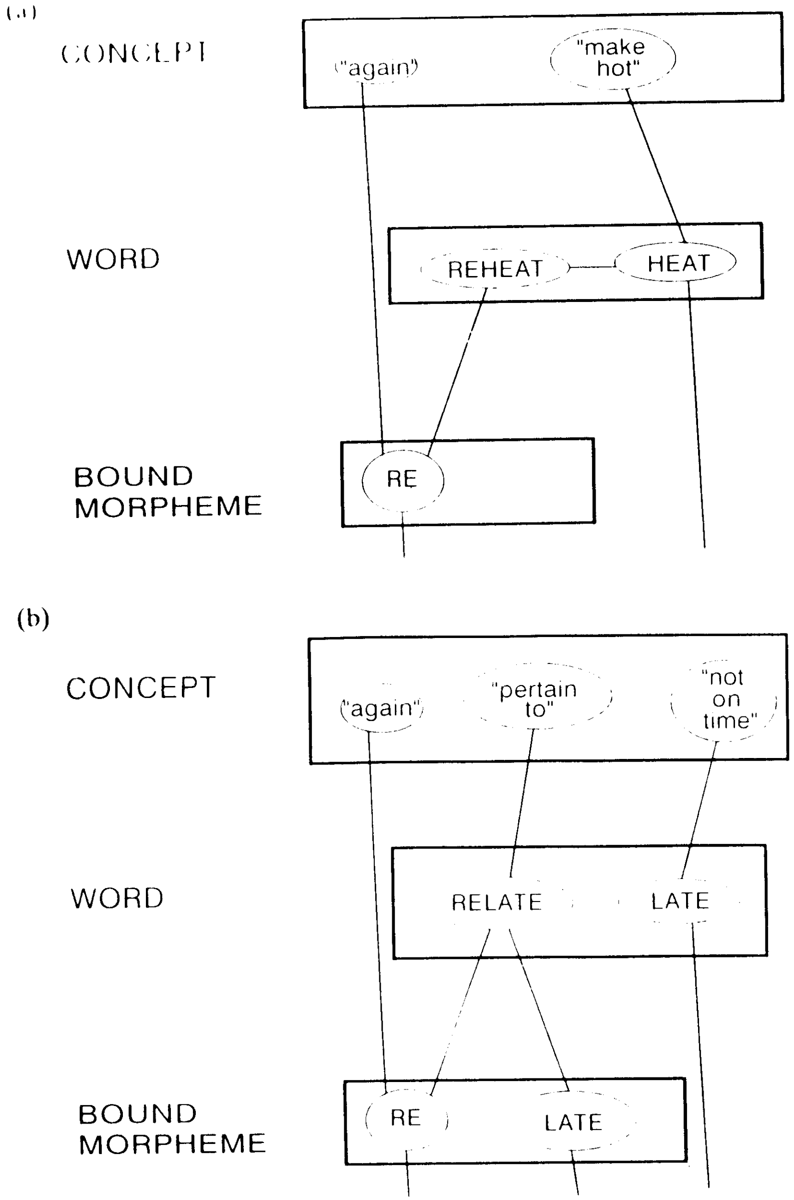


FIG. 4 (a) A possible depiction of the representation of transparent prefixed words, using REHEAT as an example. (b) A possible depiction of the representation of opaque prefixed words, using RELATE as an example.



REGRILL. Once REGRILL is encountered a few times, however, such a unit would be set up, and could even develop a connection to the concept level if it starts being used in a way that cannot be captured exactly by the combination of the meaning of its individual morphemes (as would be the case with REVIEW, for example, which conveys somewhat more than "view again").

Finally, a bound morpheme level is included rather than a general morpheme level because, otherwise, there would be no way of knowing that a free morpheme is a word while a bound morpheme (which is connected to the concept level) is not. For example, if VIVE were associated with a concept node (i.e. "live"), there would be no way of realising that it was not actually a word, since real words (like HEAT) would be represented in exactly the same way.

## Onset/Body Structure of Prefixed and Pseudoprefixed Words

The notion of a word body has in the past been largely restricted to monosyllabic words (e.g. the ENT of TENT), though here it has been extended to monosyllabic stems (e.g. the ENT of INVENT). If the body is an important unit in lexical processing, there is no reason for it to be restricted to monosyllabic words (or stems). Taft (1992) presented evidence to suggest that the first syllable of a polysyllabic word has a body representation in lexical memory (where "syllable" is orthographically defined), and has proposed that every syllable of a stem morpheme has such a representation. In this way, RELISH will have two bodies, namely EL and ISH, since it is a single morpheme. On the other hand, only the stem of REVIVE will have a body, namely IVE (or perhaps IV and E; see Taft, 1979b). Thus, the onset/body structure of prefixed words and pseudoprefixed words will usually be quite different (e.g. RE/V/IVE *vs* R/EL/ISH; EM/B/ED *vs* EMB/ER).

As pointed out earlier, such a difference in structure can explain the delay in responses to pseudoprefixed words in that the RE of RELISH will activate the RE unit that exists for the purposes of accessing words beginning with the prefix RE (such as REVIVE, REHEAT and RELATE). It should also be the case, however, that prefixed words will suffer some delay from the activation taking place in the units associated with pseudoprefixed words. For example, RELATE will activate the R and EL nodes that are appropriate for RELISH but not for RELATE. The impact of this competitive activation will presumably depend on the relative frequencies of the competing units. If a particular letter sequence forms a genuine prefix in most words in which it occurs (e.g. RE), then the prefix unit will be given considerable weight and is likely to dominate any other structures arising from that letter sequence. Hence, there will be little delay in processing a prefixed word containing that prefix. On the other hand, the O of OMIT is so unusual as a prefix that the structure OM/IT is more likely to be imposed on the word than the structure O/M/IT, and in fact, O may never even develop as a prefix unit. This conclusion appears to be consistent with recent findings (e.g. Laudanna, Burani, & Cermele, this issue) that lexical processing is influenced by the proportion of times the initial letter sequence is used in the language as a prefix as opposed to a pseudoprefix.

## Orthographic–Phonological Relationships

The focus of the discussion so far has been solely on the orthographic form of the prefixed word. However, as seen in Fig. 1, orthographic units are associated with phonological units within the interactive-activation framework. This is obviously necessary in order for visually presented words to be named. The question, though, is whether phonology is attached at the morpheme level as well as at the word level.

There must at least be a phonological representation at the word level, since some prefixed words are pronounced in a way that is not predictable from the pronunciation of their individual morphemes. For example, ASCERTAIN has stress on its final syllable, while CERTAIN does not. DEMISE has stress on its final syllable, while PROMISE does not (though it is possible that the latter is not viewed as a prefixed word at all). Without a pronunciation associated with the whole word, such irregularities would be hard to capture.

It would seem that phonology must also be attached at the morpheme level, but perhaps only for units representing prefixes. There must be a phonological representation for prefixes or else prefixed neologisms, like REGRILL, could not be pronounced. A pronunciation for RE based solely on grapheme–phoneme links would be unlikely to yield the correct pronunciation /ri:/, since E is more commonly pronounced /ɛ/ than /i:/. On the other hand, the ability to pronounce VIVE as /vaɪv/ does not necessitate a phonological representation at the bound morpheme level, since such a pronunciation could readily be generated from phonological information at the body level. It is therefore an open question whether there is a phonological representation stored for bound stem morphemes (though Taft and Zhu present evidence to suggest that there is, at least for Chinese bound morphemes). If there is not, it would seem that prefixes are represented in a qualitatively different way to stem morphemes, since only the former are linked to phonological units.



## Suffixed Words

The discussion so far has focused exclusively on prefixed words, but can the same principles be applied to suffixed words? Presumably, if a suffixed word has a free stem (e.g. HEATED, HEATER), it will be represented as in Fig. 4a, with a node representing the suffix at the morpheme level and the stem and whole word represented at the word level. It might be argued, though, that inflected forms (like HEATED) do not have their own whole word representations since they are entirely predictable from their stem, as long as information about the part of speech of the stem is also activated in some way. That is, given that HEAT can be a verb, it automatically can take on the past tense ED (unless the past tense is stored specifically because it is irregular, as in the case of CAUGHT). If this were so, in order to explain the finding that the frequency of an inflected word affects lexical decision responses even when stem frequency is controlled (Burani, Salmaso, & Caramazza, 1984; Taft, 1979a), one would have to ascribe some sort of role to the frequency with which the inflected word is generated by the addition of the suffix to the stem.

For suffixed words with bound stems (e.g. NAVAL, DENTIST, DREGS), the representation will presumably be similar to that depicted in Fig. 4b. For example, there will be units at the bound morpheme level representing NAV and AL whose combined activation will activate the word unit NAVAL. The unit NAV might in itself be linked to a concept unit for "boat-related things".

Given the suggestion that suffixed words would be represented in essentially the same form as prefixed words, how can we explain differences that have been found in the processing of suffixed and prefixed words, as well as between the processing of derivationally suffixed and inflectionally suffixed words? Taft (1985) outlines a number of experiments in English which support the idea that inflected words and prefixed words show similar effects in relation to stem and surface frequency, pseudoaffixation and nonword interference, while derived words behave like any other polysyllabic words except for stem-frequency effects (e.g. Bradley, 1979; Burani & Caramazza, 1987). The conclusion can be drawn from this that inflections are stripped off in the same way as are prefixes, while derivational suffixes are not. The stem of a derived word is stored as a unit of representation, but is accessed via a left-to-right reiterative parsing mechanism which circumvents the need to strip off the suffix prior to access.

Incorporating such notions into the interactive-activation framework is not so easy. Inflected and derived words should be represented with a similar structure (e.g. H/EAT/ED, H/EAT/ER, N/AV/AL) and, therefore, it is difficult to see why they should differ in their processing. Perhaps the difference lies in the extremely high frequency of inflectional suffixes relative to most derivational suffixes, which renders the former less informative. Another possibility is that inflections are not given a representation in the lexical system at all, but are instead processed in a separate syntactic module. They are automatically stripped off when an inflected word is presented (and inappropriately so when a pseudo-inflected word is presented, like LENS or WHITING), so that they can be independently fed into this syntactic module. If so, it would need to be concluded that the similarity of the pattern of results obtained with prefixed and inflected words is a superficial one, since the two would be processed quite differently.

The difference between prefixed and derived words must also be captured in the model, such that the prefix is ascribed the equivalent of a stripping procedure while the suffix is not. What one might say is that left-to-right processing takes place for all words, but that the morpheme units activated by the stems of derived words provide activation to the whole word level that is more effective than that provided by the morpheme units activated by the prefixes of prefixed words. For prefixed words, it is the subsequently processed stem morpheme that provides the most useful information. Problematically, however, French research has been inconsistent with the English research when it comes to a comparison of prefixed and derivationally suffixed words. Results obtained by Colé, Beauvillain, Pavard and Segui (1986) with pseudoaffixed words, and Colé, Beauvillain and Segui (1989) manipulating stem frequency, suggested that prefixed words are not morphologically analysed, whereas suffixed words are. On the other hand, Holmes and O'Regan (1992) and Beauvillain (this issue) found evidence to suggest that both types of affixed words are morphologically processed in French.

All in all, one can say that there needs to be more research in French, English and other languages to determine the relationship between the processing of prefixed and suffixed words, since it is currently unclear what the true picture is.

## Priming Studies

In addition to the research described so far, examination of morphological processes in reading has often been made using a priming paradigm. Here, an examination is made to see whether response times to a word are influenced by the prior exposure to a morphologically related form of that word. For example, Stanners, Neiser and Painton (1979) showed that response times to HAPPY were facilitated to the same degree whether primed by UNHAPPY or by HAPPY itself. It is important in the priming task, though, to ensure that responses cannot be based on an episodic



memory trace of the prime, since they would then be reflecting non-lexical strategies. When the opportunity for such episodic strategies is reduced (e.g. Forster, Davis, Schoknecht, & Carter, 1987; Fowler et al., 1985; Grainger et al., 1991), one still finds strong priming between morphologically related words.

The interactive-activation model can handle priming effects by supposing that when a node is re-used shortly after it has been activated, the lingering activation provides an advantage compared with a node which has not been recently activated. Such a view was put forward by Fowler et al., but they also point out that it is problematical for the model that facilitatory effects seem to last across quite considerable lags between the prime and target, since this seems to weaken the idea of a simple maintenance of activation within a node. Instead, it seems that there may be non-lexical strategic factors at play and, for this reason, one needs to be very cautious when interpreting the results of priming studies.

# CONCLUSIONS

What I have attempted to do in this paper is, first, to provide a description of the principal findings on which the traditional characterisation of prefix-stripping was based, presenting in the process a further experiment which improves upon one of the most important of the original studies. The suggestion by Taft and Forster (1975) was that stem morphemes form access codes which are searched for after the prefix has been stripped off. Problems with this approach were presented, however, focusing primarily on the requirement for a special independently accessible prefix store. The preferred alternative position was seen as being an interactive-activation model where prefixed words are stored in a decomposed form, but which does not require pre-lexical prefix stripping. Such a model is extremely flexible and allows us to incorporate notions about semantic transparency and shades of prefixedness.

The major alternative model currently being advocated appears to be the Augmented Addressed Morphology (AAM) model (see, e.g. Caramazza, Laudanna, & Romani, 1988; Laudanna & Burani, 1985; Laudanna et al., this issue). This model is also extremely flexible, the stem morphene being able to be accessed through two different routes: a whole word route and a stem route. The challenge, then, for the AAM model is to establish when the two different access routes come into play, and this is what advocates of this model are attempting to do (see Burani & Laudanna, 1992). For example, Laudanna et al. (this issue) propose that the more frequently an initial letter combination occurs as a prefix, the more likely a word beginning with that letter combination will be decomposed and accessed through its stem. As was explained earlier, however, the interactive-activation model comes to the same conclusion.

Whether there is any way of differentiating the performance of the AAM model and the interactive-activation model remains to be seen. Both include a representation of the whole polymorphemic word and of its morphemes, but in the interactive-activation model, activation passes through the latter to get to the former, while the two types of representation are processed independently in the AAM model. One might suppose from this that one way to differentiate the models would be to examine the interaction of word frequency with morpheme frequency. According to the AAM model, one might not witness any effects of morpheme frequency when the word is of high frequency, since whole word access might succeed before the effects of stem access come into play. On the other hand, should morpheme frequency effects be found with high-frequency words, it may be possible to ascribe these effects to the morphologically decomposed lexical entries accessed via the whole word. In the interactive-activation model, all words are accessed via their morphemes, and therefore one might predict that morpheme frequency effects will be observed for high-frequency words just as much as for low-frequency words. However, it is possible to say that very little activation is required at the morpheme level for a high-frequency word level unit to be activated, such that no effects of frequency at the morpheme level will be detected. So we see that both models appear to be able to handle any frequency data that are obtained.

Where the AAM and interactive-activation models do appear to make different predictions is in relation to the pseudoprefix effect. In the AAM model, whole word access and stem access are racing with each other, but not inhibiting each other. Since the stem of a pseudoprefixed word will not be found in the stem access system, a pseudoprefixed word should be recognised through the whole word access system in exactly the same way as a non-prefixed word. The fact that there appears to be a delay in responding to pseudoprefixed words might therefore be seen as evidence favouring the interactive-activation account over the AAM model.

While it remains to be seen what the best way is to characterise morphological processing in reading, the interactive-activation approach does seem to offer a very useful way of thinking about the mechanisms involved.

Manuscript received 21 January 1993
Revised manuscript received 30 September 1993



## REFERENCES

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

Bentin, S., & Feldman, L.B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology*, *42A*, 693–711.

Bergman, M.W., Hudson, P.T.W., & Eling, P.A.T.M. (1988). How simple complex words can be: Morphological processing and word representations. *Quarterly Journal of Experimental Psychology*, *40A*, 41–72.

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

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes*, *2*, 217–227.

Burani, C., & Laudanna, A. (1992). Units of representation for derived words in the lexicon. In R. Frost & L. Katz (Eds), *Orthography, phonology, morphology, and meaning*. Amsterdam: North-Holland.

Burani, C., Salmaso, D., & Caramazza, A. (1984). Morphological structure and lexical access. *Visible Language*, *18*, 342–352.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*, 297–332.

Colé, P., Beauvillain, C., Pavard, B., & Segui, J. (1986). Organisation morphologique et acces au lexique. *L'Année Psychologique*, *86*, 349–365.

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.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI*. New York: Academic Press.

Dell, G.S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.

Forster, K.I., Davis, C., Schoknecht, R., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology*, *39A*, 211–251.

Fowler, C.A., Napps, S.E., & Feldman, L. (1985). Relations between regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory and Cognition*, *13*, 241–255.

Grainger, J., Colé, P., & Segui, J. (1991). Masked morphological priming in visual word recognition. *Journal of Memory and Language*, *30*, 370–384.

Henderson, L. (1985). Towards a psychology of morphemes. In A.W. Ellis (Ed.), *Progress in the psychology of language*, Vol. 1. London: Lawrence Erlbaum Associates Ltd.

Henderson, L. (1986). From morph to morpheme: The psychologist gaily trips where the linguist has trodden. In G. Augst (Ed.), *International research in graphemics and orthography*. Berlin: De Gruyter.

Holmes, V.M., & O'Regan, J.K. (1992). Reading derivationally affixed French words. *Language and Cognitive Processes*, *7*, 163–192.

Laudanna, A., & Burani, C. (1985). Address mechanisms to decomposed lexical entries. *Linguistics*, *23*, 775–792.

Lima, S.D. (1987). Morphological analysis in sentence reading. *Journal of Memory and Language*, *26*, 84–99.

Marslen-Wilson, W., Tyler, L.K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, *101*, 3–33.

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*, 13–32.

McClelland, J.L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII*. London: Lawrence Erlbaum Associates Ltd.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.

Sandra, D. (1990). On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *Quarterly Journal of Experimental Psychology*, *42*, 529–567.

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

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

Taft, M. (1979b). Lexical access via an orthographic code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behavior*, *18*, 21–39.

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. Besner, T.G. Waller, & G.E. MacKinnon (Eds), *Reading research: Advances in theory and practice*, Vol. V. New York: Academic Press.

Taft, M. (1988). A morphological decomposition model of lexical representation. *Linguistics*, *26*, 657–667.

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

Taft, M. (1992). The body of the BOSS: Subsyllabic units in the lexical processing of polysyllabic words. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1004–1014.

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., & Zhu, X. (1994). The representation of bound morphemes in the lexicon: A Chinese study. In L. Feldman (Ed.), *Morphological aspects of language processing*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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.