Morphological Complexity in English Prefixed Words: An Experimental Investigation

by

Alissa Melinger

A dissertation submitted to the
Faculty of the Graduate School of State
University of New York at Buffalo
In partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

Department of Linguistics

Acknowledgments

I owe a great many thanks to the people who befriended and supported me throughout my time in graduate school. I feel as though I was extremely fortunate to have a solid support unit. It's so easy to have your self-esteem shattered in the context of a graduate program. With the possibility of rejection, criticism and failure around every bend, finishing without hating yourself or your field is rare. I attribute my semi-intact self-esteem to the praises and kind words from Karin Michelson and Jean-Pierre Koenig. Without their support, I don't know if I would have had the fortitude to continue. I also want to acknowledge the support I received from the Center for Cognitive Science and the department of Linguistics. Having financial support smoothes all bumps.

Thanks to Karin Michelson and Gail Mauner for inspiring me to pursue this topic for my dissertation. If not for the wonderful lectures and discussions in morphology class, which coincided with my training in Gail's lab, it would never have occurred to me investigate these issues experimentally. I thank Paul Luce for all his help with statistical analyses, for being so patient throughout all my numerous questions and for never mentioning the experiment that never happened. And of course I must thank Gail for inviting me into her lab in the first place and providing me with a stocked 'tool kit' for future research. I especially thank J.P. Koenig for all his advising. J. P. was the kind of advisor everyone should have. He gave me a lot of attention, albeit divided, and time to talk over my ideas and plans. He was always interested and enthusiastic and would always push me to find the really interesting issues and the big picture. So much of my intellectual growth and confidence came from him.

Serving a different but equally important role are my friends and colleagues who I used as sounding boards, proof readers, and tension release mechanisms. A special thanks goes to Wendy Baldwin for her wisdom, common sense and proof reading skills. I want to thank Eve

Abstract

This dissertation presents an experimental investigation of the representations of morphologically complex words with the intended goal of revealing how morphological information is realized in the mental lexicon. It focuses on the representations of prefixed words in English. Five segment shifting experiments (Feldman and Fowler 1987) and one speech error elicitation experiment (Baars and Motley 1976) are discussed. The results of these studies suggest that lexical representations of some prefixed words include associative links to their component morphemes while others include only weak associate links. Behavioral evidence for morphological complexity was dependent on the amount of linguistic evidence supporting complexity. Three sources of evidence which can contribute to the identification and representation of internal morphological structure were investigated: (1) whether a word is derived from a free or bound root, (2) whether the root participates in a phonological alternation, and (3) whether the prefix contributes a clear semantic component to the meaning of the whole word. Both free and bound root words provided evidence of complexity in that they elicited significantly different response times in the segment shifting experiments than phonologically controlled morphologically simple words. In the speech error elicitation study, they produced significantly more speech errors involving morphemes than monomorphemic words produced speech errors involved syllables. Evidence suggesting that the phonological form of the root is a strong cue for morphological decomposition was weak. Furthermore, results did not support the hypothesis that idiosyncratic phonological alternations, such as the alternation between *receive* and *reception* aid in the identification of a root morpheme. Finally, semantic transparency of a prefix, which was found to be a strong cue for morphological complexity, was also an interacting factor with root type. Bound root words with semantically transparent prefixes produced a different pattern of results than free stem words with semantically transparent prefixes; however,

both were significantly different from a monomorphemic control condition. In contrast, bound

root and free stem words with semantically opaque prefixes were not significantly different from

the monomorphemic control condition.

Chapter 1

INTRODUCTION

The average speaker of English knows between 10,000 and 40,000 words, possibly more depending on how *word* is defined. For each of these 10,000+ words, a speaker minimally must know how to pronounce it, what it means, and how to use it in a sentence. Not only is that a lot of information to keep track of, new words are regularly added to the existing bank of words. Speakers also know when two words are related to one another or constructed from similar parts. For examples, speakers know that *dogs* and *dog* are related because they are very similar in what they mean and how they sound. The knowledge of when words have common parts or when words sound and mean similar things is used by the speaker to organize her knowledge of words. The knowledge that words are composed of smaller parts and the knowledge of how these parts combine is called **morphological** knowledge.

The component parts of words are called **morphemes**. Morphemes are the smallest meaning bearing unit in a language. Thus, the word *unfreezable* consists of three morphemes because it can be reduced to three identifiable units of meaning — {un-}, {freeze}, and {-able}. Each of these units also occurs in other words, such as in examples (1)-(3). For example, each word in (1) begins with the phonetic string [ʌn] and each word means something like "not X", where X is the portion of the word that remains if the initial string were removed. Since these words exhibit a common phonetic string and a common meaning component, the shared string is identified as a morpheme.[1]

(1) unhappy, uncover, uninvolved, undone, unviable, unwashed, unwelcome
(2) freeze, freezing, freezable, deep-freeze, freezingly
(3) comfortable, debatable, lovable, payable, taxable

---

[1] In this dissertation, I will represent morphemes in curly brackets, {}, phonetic strings in square brackets, [ ], meaning components and definitions in quotes and word examples in the text with italics. Definitions will either be intuition based or taken from Webster's Collegiate Dictionary, Tenth Edition.

The goal of theories of morphology is to characterize the knowledge that speakers possess about the morphological structure and relatedness of words as well as to define the types of word formations which occur cross-linguistically. Knowledge of the morphology of a language allows a speaker to understand and create new words composed of familiar parts. For example, even if one has never heard the word *unfreezable*, its meaning can be inferred from the meaning and functions of its component morphemes: *freeze* is a verb which mean *"*to become congealed into ice by cold". The suffix {-able} changes this verb into an adjective meaning "able to be frozen". The prefix {un-} negates the adjective, creating a word which means "not able to be frozen". It is likely that the knowledge of the functions and meanings of these word parts is drawn from the knowledge of the many other words which contain them, such as those in examples (1)-(3). The same knowledge that allows speakers to understand new words like *unfreezable* can also restrict them from constructing unnecessary redundant words, such as **opacness* or **darkify* when *opacity* and *darken* are already known. It also provides speakers with the knowledge that **darkable *happiment* and **menthappy* are not grammatical words in English.

In addition to aiding in the production and comprehension of novel words, morphological information may also be used in the storage, production and recognition of words in the mental lexicon. Word representations in memory may be organized by shared morphemes. Also, the presentation of one word with a particular morpheme may affect the processing of subsequently presented words if they contain the same morpheme.

This dissertation investigates the representations of words composed of more than one morpheme with the intended goal of revealing how morphological information is realized in the mental lexicon. Specifically, I focus on the representations of prefixed words in English. I present evidence that the lexical representations of some prefixed words include associative links to the lexical representations of their component morphemes while others do not. I investigate

three factors that may contribute to the identification and representation of word components: whether a word is derived from a free stem or a bound root, whether the root has a phonological alternation, and whether the prefix contributes a clear semantic component to the meaning of the whole word. (These terms will be defined in the following section.) Evidence for the associative links between a word and its morphemes is supported for prefixed words derived from both bound roots and free stems when sufficient semantic content is shared between the word and its component morphemes.

1.1 The Morpheme

As mentioned above, the morpheme is defined as the smallest meaning-bearing unit. It is a minimal unit of meaning because it cannot be further broken down into smaller meaningful units. For example, {freeze} can be divided in a number of way, e.g., *fr-eeze, free-ze, fre-eze*, but none of these divisions produces two meaningful units. Further division only results in phonological units, devoid of predictable or consistent meaning.

The two characteristics that most typically indicate the presence of a common morpheme across words are similar meanings and a common phonological string. Words that only share a phonological form, such as those in (4), or only share meaning components, such as those in (5), do not necessarily contain a common morpheme.

(4) a. under, uncle
     b. table, cable, stable
(5) a. run, jog
     b. speak, say, talk, tell
     c. couch, sofa

In addition to capturing a relationship between words like *happy* and *unhappy*, morphemes have been argued to allow for a more economical storage system. The repository for words and/or morphemes in memory is called the **lexicon**. For many years it was thought that there was a premium on storage space in the lexicon. Economy of storage was highly valued and thus,

predictable information, such as the consistent form and meaning between two related words, was absent from the lexicon. Redundant full word representations such as *happy, unhappy, happiness, happily, unhappiness, unhappily*, etc, were replaced with economical morpheme representations, namely {happy}, {un-}, {-ness}, and {-ly}, which can be combined to create predictable complex words.

Morphemes also serve to capture significant linguistic generalizations and patterns. The generalization that many singular nouns in English have a plural form that ends in [-s] is lost if a plural morpheme is not posited. These generalizations can also serve as the bases for new word creation and understanding. If a newly created invention is called a *sneed,* knowledge of English morphology will allow a speaker to create the plural form *sneeds*. Morphological knowledge also describes the types of morpheme combinations that are allowable in a language. It explains how speakers know that combinations of morphemes such as **tionmovity* and **happyity* do not exist, though they may not be able to explain why they do not exist.

1.2 Definitions

Every word consists of at least one morpheme and many morphemes can be words. Morphemes that can stand alone as words, such as in (6), are called **free** morphemes. Morphemes that cannot stand alone, such as the affixes in (7), are called **bound** morphemes. Free morphemes are usually content morphemes, corresponding to concepts denoted by the major lexical categories of nouns, verbs and adjectives. Bound morphemes can be content or function morphemes. The affixes underlined in (8) and (9) are **derivational** affixes. They often change the part of speech of the word, as in (9), or modify the word such that the meaning of the whole word is different from the meaning of the base, as in (8) and (9). The affixes underlined in

(10) are **inflectional**. They generally serve a grammatical function without changing the basic meaning or part of speech of the word.

(6) happy, tree, swim, ready
(7) {un-}, {pre-}, {dis-}, {-ment}, {-ity}, {-ing}, {-ed}
(8) redo (do again), <u>un</u>do (become not done), do<u>able</u> (able to be done)
(9) govern_v, govern<u>or</u>_n, govern<u>ment</u>_n, govern<u>able</u>_{adj}
(10) jump<u>s</u>, jump<u>ing</u>, jump<u>ed</u>, jump<u>er</u>

Words composed of a single morpheme, such as in (6), are called **monomorphemic** or morphologically **simple** words. Words composed of more than one morpheme, such as in (8)-(10), are called **polymorphemic** or morphologically **complex**. Usually a polymorphemic word has a single content morpheme, which contributes the basic meaning of the word, and one or more affixes, which modify the basic meaning. When two words share a (content) morpheme, they are said to be **morphologically related.**[2]

Many languages do not have free morphemes. Even the morphemes that correspond to words of the major lexical categories (i.e., nouns, verbs, adjectives) cannot stand alone. This is commonly found in languages with robust inflectional (and derivational) morphological systems. For these languages, a minimal word will be composed of a bound content morpheme and an inflectional morpheme, such as the examples from Italian in (11). The content morpheme in (11a) is not an allowable word in Italian; it needs an inflectional affix to license it as an actual word. Examples (11b)-(11f) show that the content morpheme in (11a) contributes the core meaning of the word.

(11) a.   {lav-}   *wash(v)*
     b.   lavo   *I wash*
     c.   laviamo   *we wash*
     d.   lavare   *to wash*
     e.   lavato   *washed*
     f.   lavando   *washing*

---

[2] As will be discussed below, words that are morphologically related need not be both phonologically and semantically similar. For example, while the words *permit* and *remit* are argued to be morphologically related, they are not **semantically related**. Similarly, while the words *go* and *went* are not **phonologically related**, they are still morphologically related.

Since (11a) cannot stand alone as a word, it is not a free morpheme. However, most of the bound morphemes discussed thus far have been affixes and (11a) is not an affix. Thus another classificatory distinction must be made between **roots**, **stems** and **bases**.[3] **Roots** are content morphemes that cannot be further divided into smaller morphemic parts. A root need not be a free standing word. Thus, both {lav-} in *lavando* and {burden} in *unburden* are roots; the latter being a free root and the former being a bound root. A **base** is anything that an affix attaches to. For example, in the word *unlovable*, *love* is the base for *lovable*, and *lovable* is the base for *unlovable*. A **stem** is a base to which inflectional affixes attach. Thus, {lav-} and {burden} are at the same time roots, stems, and bases. They are roots because they cannot be reduced to any smaller morphemic units. They are stems because inflectional affixes can be directly concatenated to them. They are bases because complex words can be derived from them by adding additional affixes. In this dissertation, I will be contrasting prefixed words like *unhappy* and *repaint* to prefixed words like *receive* and *deflect*. The former are derived from free bases, {happy} and {paint} respectively, while the latter are derived from bound bases, {-ceive} and {-flect} respectively. Strictly speaking, all of these morphemes are roots since they all take derivational prefixes. They are all bases since larger words are built off of them. However, the free morphemes {happy} and {paint} are also stems since inflectional affixes can attach to them directly, forming an allowable word in English. Thus, to avoid confusion between the sets of words I investigate, I will refer to words with free bases as **free stem words** and to words with bound bases as **bound root words**.

One characteristic of bound root words in English is that they often do not have an easily perceived common meaning component. Consider the words in (12).

---

[3]Authors use these term differently. The definitions proposed here are not intended to supplant prior usages but rather lay out the way I will be using them in this dissertation.

(12)   concede   commit   consist   conceive   conflate
         recede   remit   resist   receive
         secede   submit   subsist
                     permit   persist   perceive
                                desist   deceive   deflate
                                insist   inflate

In (12) we have rows and columns of words organized by what look to be common morphemes. Each row appears to share a prefix and each column appears to share a root. However, the words in each column do not have a common meaning component that is easily apparent to most native speakers. The words in (13) might also look as though they are composed of prefixes and bound roots.

(13) religion, regard, perish, constant, denial,

For example, it is possible that *religion* is composed of the prefix {re-} and the base {-ligion}; *constant* could be composed of the prefix {con-} and the base {-stant}. This analysis is generally rejected since the strings that are identified by such an analysis, namely {-stant} and {-ligion}, are not related to the meaning of the words in which they occur and do not occur in other words. Note that there is no word *prestant, *destant,* or *exstant.* In contrast, the words in example (12) are more plausibly analyzed as polymorphemic since their roots occur in many words.

Additional evidence for the complexity of the items in (12) comes from the fact that they participate in idiosyncratic phonological alternations when they undergo suffixation, such as in (14).

(14)   secede ~ secession   submit ~ submission
         recede ~ recession   remit ~ remission   receive ~ receptive
         concede ~ concession   commit ~ commission   conceive ~ conceptive
                                  permit ~ permission   perceive ~ perceptive
                                                          deceive ~ deceptive

The phonological alternations in (14) are not phonetically conditioned nor are they productive within the language. Other words which have similar word endings do not alternate, and they select other forms of the suffixes or a different suffix altogether, such as in (15).

(15) a. limit ~ limitation, * limission
b. vomit ~ vomitory, * vomission
c. retrieve ~ retrieval, * retreption

Why should the words in each column of (14) undergo the same unproductive phonological alternation while the words in (15) do not? How does the phonological system discriminate between words that undergo the alternation and those that do not? The claim made by some linguists is that idiosyncratic phonological rules apply only to specific morphemes. Thus, if two words share the same word ending and undergo the same phonological alternation, then they must share a morpheme.

Characterizing the columns of words in (12) and (14) as sharing a root explains why they participate in unusual phonological alternations, thus capturing a significant linguistic generalization. But it also serves to complicate the way a morpheme is defined. Note that these roots and other bound roots in English tend not to contribute a consistent or clear meaning to the whole word. I will refer to morphemes that do not provide a consistent or recognizable meaning component to a word as **semantically opaque**. In contrast, morphemes that provide clear and consistent meaning will be referred to as **semantically transparent**.[4]

The definition of Morpheme as the smallest unit of meaning does not allow for the classification of English bound roots as morphemes. The failure of the above definition to capture English bound roots as morphemes is by no means a new observation and several linguists have addressed it in the literature (cf., Nida 1946, Hockett 1954, Jackendoff 1975,

---

[4] Note that semantic transparency and morphological transparency are not equivalent. Semantically opaque morphemes are often easily identifiable as morphemes despite the fact that their meaning is unclear. Similarly, words whose meanings are not clearly related to the meanings of their component parts can still be identified as composed of subparts. In this dissertation, the terms "transparent" and "opaque" refer to the semantic contribution of the morphemes, not the saliency of the individual morphemes, unless stated otherwise.

Aronoff 1976). In fact, there are other types of morphemes that challenge this simplistic

definition. One type of morpheme that challenges this definition is the **cranberry morpheme**.

Cranberry morphemes are bound morphemes that occur only in a single word. Their meaning is

derivable from whatever word semantics remain after the contributions of the other morphemes

have been considered. For example, the meaning of the morpheme {cran-} is whatever meaning

the word *cranberry* has that is not derivable from the meaning of *berry*. Another problem for the

definition of the morpheme is that some morphemes can be realized without any phonetic form.

These null realizations of a morpheme are called **zero morphs**. For example, the plural

morpheme in English can be realized as a suffix, as in *dish ~ dishes* or *ox ~ oxen* or it can have

no phonetic realization, as in *sheep$_{sg}$ ~ sheep$_{pl}$*. Although the form of *sheep* does not change, it

can still express the meaning of the plural morpheme. Cranberry morphemes and zero morphs

illustrate how the simplistic definition of the morpheme as the smallest unit of meaning falls

short. Cranberry morphemes are not meaningful independent of the single word in which they

occur and zero morphs are not units of form.

   To address these problems with the definition of morpheme, Aronoff (1976) argued that a

morpheme should not be defined as the smallest meaningful unit. Rather, he suggested that a

morpheme be defined as a "phonetic string which can be connected to a linguistic entity outside

that string. What is important is not its meaning, but its arbitrariness" (1976:15). The linguistic

entity can be, and often is, an arbitrary but consistent semantic content. But it can also be a

phonological process or a grammatical function such as deriving a plural form from a singular

form. In the absence of shared semantics, the presence of a formal alternation can serve to

distinguish a string as a root morpheme. Thus, the alternating words in each column of (14) have

a common root while the words in (15a)-(15b) do not.

Irrespective of how Morpheme is defined, it is acknowledged that the meanings of morphemes are arbitrary. In contrast, when morphemes are combined to produce words, the meaning of these words are often predictable from the meanings of their components. When the meaning of a complex word is fully predictable from the meanings of the component morphemes it is called **semantically compositional**. The words in (16) are **fully compositional**.

(16) a.   dogs
     b.   unhappy
     c.   actor
     d.   singable

Each of the words in (16) can be defined as the sum of the meanings of their parts. In (16a), the meaning of *dogs* is the meaning of {dog} plus the meaning of the plural suffix. Likewise, in (16b) the meaning of *unhappy* is fully predictable from the meaning of the prefix {un-} and the meaning of the root {happy}.

While this is the ideal situation, it is not uncommon for words to be less than fully compositional. Phenomena like semantic drift, metonymy, and lexicalization have created, many words which no longer express a fully compositional meaning, even when they consist of semantically transparent morphemes. Consider the examples in (17). While the meaning contribution of each individual morpheme to the meaning of the whole word can be seen rather easily, the meaning of the whole word additionally includes some idiosyncratic meaning component not provided by the individual morphemes.

(17) a.   subdivision
     b.   computer
     c.   redcoat

Example (17a) has (at least) two meanings; one is fairly compositional, meaning "an act or instance of dividing into parts". The second meaning is a "tract of land surveyed and divided into lots for purposes of sale". In both cases the contribution of the individual morphemes can be seen, but in the latter definition there is an additional sense that is only associated with the

meaning of the whole word, namely the reference to the tract of land. Similarly for (17b), a

computer is much more than just someone or something that computes. And a *redcoat,* in (17c),

does not refer to an article of clothing. In all of these words there is a portion of the word's

meaning that is not retrievable from the components.

   The transition from total compositionality to the complete absence of compositionality is

gradual. Many words are partially compositional, consisting of some transparent and some

opaque morphemes, as in the examples in (18).

(18) a. retrieve
      b. contribute
      c. shoehorn

For many speakers, example (18a) has a somewhat transparent prefix while the root's meaning is

inaccessible. The meaning of the prefix {re-} is "back" or "again". The meaning of *retrieve* is

"to get back again", according to Webster's Collegiate Dictionary. The semantic contribution of

the prefix {re-} to the word *retrieve* is clear from this definition. However, the meaning of the

root, {-trieve}, is opaque. Example (18b) illustrates a word with a semantically transparent root

but a semantically opaque prefix. While many speakers may define *contribute* as "to give a

tribute*"*, it is not clear what meaning the prefix {con-}, which is usually defined as "together,

with", provides to the meaning of the word. Thus, despite the fact that the prefix {con-} occurs

in more words than the root {tribute} and, in turn, may be identified as a unit by speakers, its

meaning is opaque. In example (18c), the contribution of the first morpheme, *shoe*, is relevant to

the meaning of the whole word, but the relationship between *horn* and *shoehorn* is unclear to

most speakers.

Words can also be totally fossilized, such as the words in (19). In these cases, the semantic contributions of the component parts of the words have become completely lost; the meanings of the words are completely unexpected given the meanings of the component parts.[5]

(19) a.   pitcher (of beer)
     b.   transmission
     c.   receive

Words in (19) have meanings that are only remotely related to the meanings of the individual morphemes, if at all. For example, the word *pitcher* in (19a) can refer to a container or to a baseball player. When it refers to a container, it is not related to the meaning of its parts at all; its meaning is completely idiomatic. In fact, some speakers might not identify this word as complex at all, despite the surface evidence for complexity. Example (19b) is, for most people, also removed from the meaning of its component morphemes. The compositional meaning of the word refers to a message while the non-compositional meaning refers to a part of an automobile. For many people, these two words, *transmission₁* and *transmission₂,* are separate homophones not unlike the separate meanings of the word *duck.*

At this point I have presented some arguments for why morphemes are needed, developed some terminology for talking about different types of morphemes, and identified different types of evidence for morphemes. I have also discussed how morphemes can combine to produce words whose meanings are fully compositional, partly compositional, or non-compositional. I will turn now to a brief discussion of some of the competing models that have been proposed for the representation of morphological information and discuss their implications for lexical organization.

---

[5] It should be mentioned that even words that have become fully fossilized usually have another (sometimes secondary) meaning which is fully compositional, as we saw in example (18a).

1.3 Morphological Theories

Many models have been developed specifically to account for a particular set of problematic data. Since it is beyond the scope of this dissertation to summarize all the past theories, I will try to summarize some of the more influential theories that have been proposed for data most relevant to the types of morphological combinations discussed in this thesis. The experiments presented below may not serve to distinguish between specific instantiations of these models and the distinctions made between theories may not bear on the interpretation of my results. However, it is still useful to have some background on the theories to frame the upcoming discussion, since the issues addressed in these theories motivated some of the investigations to be presented below.

As is often the case, a discussion of competing theories is often best understood as a set of contrasts. The first contrast is between **word-based theories** and **morpheme- based theories.** The primary distinction between these theories is in what they consider to be the essential units, or building blocks, for word formation.

For morpheme-based theorists, morphemes are the building blocks for word creation; morphemes are combined, via concatenation, to create the words of the language. Furthermore, most morpheme-based theories assume that roots and affixes both have their own independent lexical representations. Some morpheme-based theories additionally list (some) complex words in the lexicon, but this is not a crucial assumption of these theories (Halle 1973, Lieber 1980, 1992, Selkirk 1982). For word-based theories, words are the building blocks of word formation processes (Jackendoff 1975, Aronoff 1976, Bochner 1993), not morphemes. As a result of this restriction on what can be input to a word formation process, most word-based models assume that only real words are listed in the lexicon; bound morphemes (i.e., bound roots and affixes) do not have a lexical representation independent from the words in which they occur.

13

Another theoretical contrast is between **Item and Arrangement** models (IA models hereafter) and **Item and Process** models (IP models hereafter).[6] The crucial difference between these two models is in how they envision the process of word formation. For IA models, all morphemes are listed in the lexicon. Word formation occurs by selecting the required morphemes and concatenating them together to produce the desired complex word. For example, if a speaker wants to produce a word that is the opposite of *happy* she selects the base {happy} and the appropriate affix, in this case {un-} and concatenates them together, to produce *unhappy*. In contrast, IP models imagine that word formation takes place by applying **rules** of word formation, not via concatenation. The IP models distinguish between the representations of roots and affixes. Roots have lexical representations in the more traditional sense while affixes are represented as rules or processes which are associated with **morphs**, phonetic strings that realize the morpheme. To create the word *unhappy*, the base {happy} undergoes a word formation process to derive *unhappy*. Crucially in IP models, the prefix {un-} has no lexical representation.

For the most part, IP models are word-based, since affixes do not have any independent lexical status. However, some IP models allow for bound roots to be the input to word formation processes. Similarly, most IA models are morpheme-based (however, see Kiparsky 1982 for a discussion of what could be considered a word-based IA model in Lexical Phonology).

Now that I have outlined the relevant assumptions behind the different theories, I will introduce some particular instantiations of these models and examine what specific predictions each models makes about word relatedness. In the interest of time and topicality, I will limit my discussion to a summary of the representational assumptions these theories make for derivationally related words.

---

[6] These terms are quite old and their usage has changed in many ways over the years (see Hockett 1954 for an early discussion of the contrast).

There are several ways in which lexical relationships could be represented in a morphological model. One possibility is that words are productively produced from morphemes on-line and that morphologically complex words have no permanent memory representation. Many early theories of morphology hold this assumption, but more recent theories give complex words a permanent memory representation in the lexicon. However, many theories separate the list of morphemes from the list of complex words (e.g., Halle 1973) while others only list a subset of complex words (Anshen & Aronoff 1988).

Following the assumptions of an IA model, complex words have internal structure, as seen in (20) and (21). Some IA models propose a relatively flat word structure, as in (20) (Chomsky & Halle 1968, Halle 1973), while others build hierarchical structure into their representations, as in (21) (Lieber 1980, Selkirk 1982). Structure can be maintained in the lexical representations of complex words for those theories that store complex words in the lexicon. For models that do not allot permanent memory representation to complex words, the structure is knowledge about how the word is produced.

(20)   a.   un + happy +ness
         b.   happy + ness
         c.   sub + divide + ion
         d.   divide + ed
(21)   a. $((un(happy)_A)_A ness)_N$
         b. $((happy)_A ness)_N$
         c. $(standard)_N ize)_V s)_V$

Since many IA models, like Lieber's (1980), do not store complex words, the notion of a lexical relation is extrinsic to the model. Instead, words are morphologically related if they are both constructed out of the same root. Knowing that two words are related comes from the knowledge of which morphemes are used in the formation of the words. Additionally, many IA models provide permanent lexical representations to words derived via non-productive or irregular morphological rules. Complex words derived by irregular rules will be related to their

stems via lexical connections that are external to the actual lexical items. These connections are

statements that relate sets of words. For example, Lieber's connections, which she termed

morpholexical rules, "... state absolutely that lexical items X are related to lexical items Y."

(Lieber 1980:40). To illustrate, consider the words in (22a) and (22b). These words hold a

regular relation in that the words of (22b) are the noun forms of the verbs in (22a). Since neither

the words of (22a) or (22b) are arguably produced via a regular process of affixation, Lieber

represented them in her lexicon.[7] The relationship between each word pair is captured by a

single morpholexical rule. The morpholexical rule is independent of the words themselves,

which are unstructured morphologically.

(22) a. run_v, $jump_v$, $hammer_v$, $fish_v$, $knock_v$, $paint_v$
      b. run_n, $jump_n$, $hammer_n$, $fish_n$, $knock_n$, $paint_n$

Word-based models maintain that morphemes are not represented in the lexicon at all; only

words are represented. Furthermore, some word-based models do not include any internal

morphological structure in the representations of complex words. If internal structure is not

included, then words that share a common root are related to each other via lexical statements

that define the relationship between two words. These statements have been formalized in the

literature as morpholexical rules (see above), Word Formation Rules (WFR hereafter), such as in

(23) or Lexical Redundancy Rules (RR hereafter), as in (25).

(23) $[X]_{Adj} \rightarrow [un \# [X]_{Adj}]_{Adj}$
     semantics (roughly) un#X = not X

(24) happy ~ unhappy, lucky ~ unlucky

---

[7] Lieber's lexicon includes only irregular morphological relations. Any word formation process that is productive and regular with a compositional output is concatenated on-line, not stored. Thus, her morpholexical rules relate pairs of words that are irregularly related, such as *take ~ took*.

(25)

$$
\begin{bmatrix}
X \\
/y + \text{ion}/ \\
+N \\
+[NP_1\text{'s} \underline{\hspace{1cm}} (P)\ NP_2] \\
\text{Abstract result of act} \\
\text{of } NP_1\text{'s Z-ing } NP_2
\end{bmatrix}
\Leftrightarrow
\begin{bmatrix}
W \\
/y/ \\
+V \\
\ [NP_1 \underline{\hspace{1cm}} (P)\ NP_2] \\
NP_1\ Z\ NP_2
\end{bmatrix}
$$

(26)   cooperate ~ cooperation, designate ~ designation, *aggress ~aggression

The WFR in (23) is taken from Aronoff (1976: 63). It expresses a relationship by which adjectives are derived into complex words, relating the word pairs in (24). This WFR also serves to analyze words into their constituent parts. When a speaker hears the word *unlucky* for the first time, it is associated with this rule since its meaning and the part-of-speech of the base fit the WFR's specifications. Consequently, the structure [un + lucky] will be imposed on the lexical entry for this word. Each word in (24) has an independent, fully specified, lexical entry, but the prefix {un-} has no lexical representation. It is only associated with words by the process, or rule, described in (23).

The lexical redundancy rule in (25), taken from Jackendoff (1975:642), expresses a symmetrical relationship between X- type and W-type lexical entries. This rule reads "A lexical entry *x* having such-and-such properties is related to a lexical entry *w* having such and such properties" Jackendoff (1975:642). This particular rule would relate the word pairs in (26), all of which have independent, fully specified, lexical entries.

WFRs and RRs both can analyze a bound root word into component parts without positing a lexical entry for the root. To illustrate, words like *aggression*, *aggressive, aggressor,* are composed of the root {aggress-} and the suffixes {-ion}, {-ive} and {-or} respectively. But there is no word *aggress*. Since a word-based model allows lexical representations for words only, there is no lexical representation of the bound root {aggress-}. However, both WFRs and RRs

17

can express the relationship between the words derived from this root without the need for an explicit root representation. A word like *aggression* can not be created by a WFR, but it can be analyzed by one. A WFR that expresses the relationship [X → X+ion] analyzes a word like *aggression* into the structure consistent with the rule, namely [aggress + ion] without the need for an independent root representation for [aggress]. Likewise, in a model which employs a RR expressing the same relationship, [X ← X+ion], the word *aggression* can map to one component of the statement without requiring that any lexical item be mapped to the opposite component. Thus, this RR can identify *aggression* as matching the statement of redundancy regardless of the fact that there is no root {aggress}. Furthermore, since both of these models are word-based, affixes have no representation other than the WFR or the RR. In a sense, the rule is the affix.

One crucial difference between the WFRs proposed by Aronoff and the RRs proposed by Jackendoff is that the WFRs are active participants in the lexicon. They are processes used generatively to create and analyze new words. In contrast, RRs are relations over a (fairly) static set of words in the lexicon. They are not processes that create words but rather they are statements that express relationships between words.

While some models propose WFRs or RRs as the sole mechanism for encoding morphological information, other theories also include explicit structure within lexical entries. For example, Jackendoff's (1975) model does not acknowledge that words have internal structure. Each word is represented as a single indivisible unit in the lexicon. RRs express relations amongst words, but no structure is imposed on the word. Jackendoff claims that speaker intuitions of relatedness demonstrate only that words are related, not that they have some compositional relationship. He considers the relationships that speakers perceive, and that linguists impose on the structure of words, to be common patterns that are followed by a number of words. When enough words follow a pattern, a RR is formed to relate the words. For

Jackendoff, words can be related without the need for them to be linked or lexically connected in any way.

In contrast, Aronoff's WFRs do produce internally structured representations, whether the WFR was used to produce or analyze the word. In fact, Aronoff discusses the use of bracketed structure in distinguishing words with compositional semantics from words whose meanings have become more lexicalized. Note the contrasting bracketed structure Aronoff proposes for the compositional and non-compositional meanings of *prohibition* in (27a) and (27b) respectively (1976: 25).

(27) a. [[pro=hibit+]$_V$ +ion]$_N$ "The act of prohibiting something"
     b. [pro=hibit+tion]$_N$ "The period in the 1920's when alcohol was prohibited"

Example (27a) includes information indexing the word to the WFR that produced it; thus it is tied to the compositional semantics attributed by its base and the WFR. Example (27b) has lost its association to the WFR that created it. Its semantics is therefore less constrained (and less predictable).

Moving away from the established dichotomy of IA vs. IP models, Bybee (1988, 1995) presents a model of lexical representation, not of word formation. Bybee proposes that, as words are learned they are stored in the lexicon with phonological, semantic, and other linguistic features. Words are stored and lexical connections are developed between words which share features. The more connections there are between words, the closer their lexical relationship is perceived to be. When phonological and semantic connections run in parallel, you have the equivalent of a morphological relation.
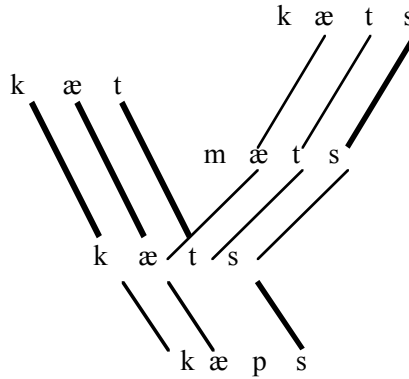
Figure 1. Bybee's connections in the lexicon

In Figure 1, taken from Bybee (1995:429), I present a subsection of the lexicon which illustrates different strengths of connections proposed in Bybee's model. For example, since the words *kæt* and *kæts* share both phonological and semantic features, they have thick lines connecting each of the segments. These thick lines, which can be viewed as shared phonological features overlaid with shared semantic features in the two- dimensional diagram, denote morphological relatedness, since phonological and semantic features are being shared in parallel. Segments that share only phonological features are connected with thin lines. The words *kæts* and *mæts* are connected too, but by lighter lines. This represents the fact that these two words, while phonologically similar, do not share semantic features. The only semantic connection between these two words is between the two final -*s*. This single dark line represents the shared semantic feature of *plural* associated with this segment. Again, the final -*s* embodies the overlap of phonological and semantic features being shared, i.e., the classic definition of a morpheme falling out as an epiphenomenon due to the function of the connections.

   In this model there is no independent morphological component; there are no explicit WFRs to derive complex words from existing ones; there are no generalizations encapsulated in RRs. Rather, there is only the lexicon and connections between shared features across the

representations. Word formation, in this model, occurs via an analogical model. For example, if the speaker wishes to produce a past tense form of a newly learned word, she searches her lexicon for the most frequent phonological string that cooccurs with the desired semantics.

Thus far I have outlined four possible representational systems for encoding morphological relations: 1) roots, affixes and irregularly derived words all have equal representational status and they are combined into complex structured words which do not have permanent memory representations (Lieber 1980), 2) affixes have no representation and relations between words are expressed as passive statements of redundancy (Jackendoff 1975), 3) affixes are assigned by rules, words are related by rules and are internally structured (Aronoff 1976), 4) whole words are represented as unstructured wholes with connections to other words in the lexicon that have similar features. Morphemes are perceived when form connections and meaning connections co-occur (Bybee 1988).

In terms of similarity, models 1) and 3) are quite similar. They differ primarily in the status given to affixes and bound roots. In contrast, models 1) and 2) are diametrically opposed, one relying solely on structured representations composed of morphemes and the other relying solely on relationships between existing words with no explicit representations for morphemes at all. Models 2) and 4) are similar in that morphemes are not explicitly included in the lexicon. But in 4), relations are expressed by connections between shared linguistic features rather than being expressed by statements of lexical patterns.

1.4 Psychological Models

In addition to linguistic theories proposing different methods for representing morphological relations, psychological models of the lexicon have also been proposed. Two such proposals that are particularly relevant to this thesis differ mainly in whether they posit that morphemes have any explicit representation. One model, which I shall refer to as the **explicit morpheme model**

(EM model hereafter), claims that words are composed of morphemes and that morphemes serve as an organizing principle in the lexicon. There are many versions of the explicit morpheme model, some mirroring the beliefs of morpheme-based linguistic theories while others echo the views of word-based models. For example, Taft (1988) proposes that all morphologically related words share a single lexical entry. Derived words are stored within the entry of their base. Another instantiation of the EM model assumes that derived words and their bases have separate representations but that they are lexically related. Some of these models assume a 'satellite' organization, where derived words are linked to their base word's representation (Lukatela et al. 1978, 1980, Feldman & Fowler 1987), while others assume that complex words are linked to their root morpheme representation (Fowler et al. 1985, Schreuder et al. 1990, Grainger et al. 1991, Schriefers et al. 1991, 1992). The crucial characteristic that all EM models share is that morphological information has an explicit and independent representation in the lexicon, either in the form of connections between related words or in independent representation of root morphemes. In this way, they are similar to most of the linguistic models presented in section 14. Different instantiations of EM models, like the linguistic models, make claims about whether the basic units of lexical relations are morphemes or words. The second class of models will be called the **implicit morpheme model** (IM model hereafter). This model states that morphemes have no independent or explicit representation in the lexicon (McClelland & Rumelhart 1985, Seidenberg & McClelland 1989, Plaut & McClelland 1993, Rueckl, Mikolinski, Raveh Miner & Mars 1997, Rueckl & Raveh 1999, Plaut & Gonnerman 2000). Morphological relations are the artifacts of phonological and semantic overlap in patterns of lexical activation (which correspond to words). The closest linguistic analog to this type of model is Bybee's model of lexical representation, in which morphemes are not explicitly represented but rather

emerge as the result of overlap in semantic and phonological connections between words. Crucially, however, Bybee's model is of representations, not processing.

Many IM models are models of word recognition rather than models of lexical representations. While they do often make claims about the processing and representations of morphological information, their primary aim is to account for all factors that have been found to influence processes of word recognition, such as frequency of the input. Stated very simply, the process of word recognition is a process of mapping an input stimulus to a memory representation. This mapping process can take place over several successive levels of representation. For example, some models of word recognition have an orthographic level of representation which is linked to a lexical level of representation which is in turn, followed by a conceptual level of representation. An orthographic input is matched to individual nodes (or clusters of nodes) at the level of orthographic representation. While the orthographic nodes are still receiving activation from the input, they begin to send activation to the lexical level of representation concomitant with sending inhibitory activation to competing orthographic representations at the same level. As representations at the lexical level are activated, they activate conceptual representations, as well as bolster the activation of the orthographic nodes at the prior level of representation and inhibiting competitors at the lexical level. Eventually, the system reaches a resonant state, meaning that the activated nodes at the different levels are mutually reinforcing each other. Once a resonant state is achieved and a single representation at the lexical level reaches its activation threshold, it is successfully recognized as the word presented in the input.

To illustrate this process, consider the following simplified example, based loosely on the model proposed by McClelland and Rumelhart (1985). Imagine the word *cat* is input to an IM model with the architecture described above. The first letter, *c*, would send strong activation to

the node (or nodes) consistent with this letter. However, it would most likely also send some activation to the nodes for the letter *o*, which is similar in appearance. The nodes for *c* and *o*, which are both receiving activation from the input, then begin to send out excitatory activation to the lexical level, activating all word forms that begin with the letters *c* and *o*, respectively, as well as sending inhibitory activation to each other. Since *c* is receiving more activation from the input (because it is a better match than *o*), it will have stronger activation and therefore be able to send out more inhibitory activation to *o* and more excitatory activation to the lexical level. At the lexical level, all words that begin with *c* or *o* have begun to receive activation. They mutually inhibit each other while feeding back excitatory activation to their corresponding orthographic nodes at the level below. Since *c* was more active than *o*, the words that begin with *c* will be more activated than the words that begin with *o*. Thus, the words with *c* send more feedback to the orthographic level than the words beginning with *o*. Eventually, all the words beginning with *o* and the orthographic nodes corresponding to *o* will be sufficiently inhibited that the only words still activated begin with *c*. The same pattern of spreading excitatory and inhibitory activation occurs for the next two letters of the input *cat* until the word is correctly identified.[8]

   This is a simplistic picture of how a word recognition model might work. I should stress that the architectures and mathematical models used by these networks vary greatly. Not all models allow for feedback from the lexical level to the orthographic level, for example. The crucial point to take away from this illustration is that words are recognized because a lexical representation, or pattern of activation, receives more evidence from the input than any other word in the network. The more clear evidence there is and the fewer competitors there are, the faster recognition occurs.

---

[8] The relative order of graphemes in the input is maintained at the orthographic level of activation. Thus, an input like *cat* will not also activate words like *act*.

Now that I have described the basic architecture of word recognition models, I will briefly outline the relationship between this architecture and IM models' interpretation of morphemic effects. Essentially, after a word has been recognized, the activated nodes take some time to return to their resting state; this is called residual activation. If another word which shares many features with the prior word is presented before the activated nodes can return to their resting state, then the nodes which are being activated afresh have a head start in reaching their activation threshold. This effect is called **priming**. When two morphologically related words are presented sequentially, they are expected to produce a priming effect since their shared phonological and semantic representational nodes are already partially activated. Priming effects will be discussed again in section 2.1. Psychological evidence supporting EM and IM models will also be presented in Chapter 2.

1.5 Goals and Focus

This dissertation focuses primarily on an investigation of the nature of the lexical representations of prefixed words in English. Specifically, I will be searching for evidence of morphological structure in the lexical representations of prefixed words. I will investigate three characteristics of prefixed words to see if they influence the way in which morphological information is lexically encoded: 1) the status of the root as a free or bound morpheme, 2) the phonological constancy of the root, and 3) the semantic transparency of the prefix. I conduct six experiments to investigate these issues using an experimental task designed to be sensitive to internal morphological structure and to the separability of an affix from the word base. The results of these experiments suggest that both free stem and bound root prefixed words can include morphological structure if sufficient linguistic evidence is present to identify the morphological components. My data will suggest that a phonological alternation (versus phonological constancy) is not a strong cue for the identification of a morpheme while the

semantic contribution of the prefix is a strong cue for complexity. These results are interpreted with respect to the linguistic models presented in section 1.3 and a discussion of which models best capture the results is presented in the conclusion. Minimally, my results suggest the need for permanent memory representations for complex words (contra Lieber 1980), and for the inclusion of graded lexical relationships between words and their component morphemes (contra Bybee 1988). Furthermore, I argue that these relations are actual links between lexical items, not expressions of redundancy (contra Jackendoff 1975).

This dissertation introduces prefix transparency and phonological constancy as factors that contribute to the inclusion of morphological structure in the lexical representations of prefixed words. While there have been many prior studies which investigate the role of root semantics in the representation of morphological information, no prior study has investigated the semantic transparency of affixes and their contribution to the overall representation of the word. Also, while the amount of phonological overlap between morphologically related words has been investigated in prior studies, no study has investigated the contribution of a phonological alternation to the identification of morphemes with a task that does not introduce both forms of the root. The results of the present experiments argue for a model of morphological information in which morphological relations can be gradient; morphological relationships need not be strictly dichotomous. Furthermore, the present results suggest that semantics plays a central role in the identification of prefixes in English, perhaps more so than for other morpheme classes.

1.5.1 Why prefixes

Prefixes in English exhibit several distinct properties from suffixes which makes their investigation interesting. These differences suggest that representational claims based on studies investigating suffixed words cannot necessarily be extended to prefixed words. One obvious difference between prefixes and suffixes is that prefixes occur before the base while suffixes

follow it. Therefore, prefixes may visually/acoustically hide the beginning of the word base, which provides the core meaning for the word. Having a prefix which masks the beginning of the base could greatly complicate word recognition. This is a fundamental claim held by some models such as the Prefix Stripping model (Taft & Forster 1975), which claims that all prefixes must be stripped away to reveal the root, the unit by which words are recognized.[9]

Another difference between prefixes and suffixes is in their influence on the base. For example, bound roots that undergo suffixation generally retain more consistent semantic content across their occurrences than prefixed bound root words, e.g., compare the {aggress-} examples, *aggression, aggressive, aggressor* to the {*-ceive*} examples, *receive, conceive, deceive*. Additionally, as mentioned in section 1.1, suffixes are associated with various formal properties such as a change in the grammatical category or phonological form of the base. Some suffixes regularly trigger phonological alternations or stress shifts, such as in (28), and others systematically change the part of speech of the root, as in (29). However, prefixes do neither in English.

(28) a. {-ity}, electric ~ electricity, vain ~ vanity, chaste ~ chastity, humid ~ humidity
     b. {-ive}, alternate ~ alternative, cooperate ~ cooperative
(29) a. {-ness} happy ~ happiness, ugly ~ ugliness, weak ~ weakness
     b. {-ment} govern ~ government, attach ~ attachment, pay ~ payment

1.5.2 Free stem words in English

Much evidence is available in the experimental literature to support a complex representation of free stem words. However, free stem words can vary in their degree of semantic compositionality, as discussed above. One particular characteristic of free stems words which is directly related to the issue of semantic compositionality will be investigated, namely the

---

[9] It is possible that prefixes present a greater complication to spoken word recognition than to visual word recognition since acoustic information is necessarily arrayed over time while many written words can be read in a single fixation.

semantic transparency of prefixes. Free stem words with semantically opaque prefixes will be shown to exhibit weaker traits of morphological complexity than free stem words with semantically transparent prefixes.

1.5.3 Bound root words in English

One major difference between EM and IM models of the mental lexicon is whether morphemes are considered explicit lexical entities or whether they emerge from phonological and semantic sources of information. This issue is sometimes discussed as the "psychological reality" of morphemes. The two types of models make different predictions regarding whether the prefixed bound root words I investigate exhibit morphologically complex traits. EM models give independent status to morphemic information and therefore predicts that evidence for the morphological complexity of bound root words should be obtainable. In contrast, IM models predict that words should only exhibit traits of morphological complexity if they also have semantic overlap with morphological relatives. The morphological relationship between bound root words, such as *receive* and *deceive*, often escapes native speakers' awareness since they lack a clear semantic relationship. In contrast, the morphological relationship between *governor* and *government* or *preheat* and *reheat* are more obvious to speakers since these word pairs share a strong meaning component. Thus, IM models predict that words will not exhibit morphologically complex behaviors in the absence of strong evidence from semantics and phonological form. This is especially true in a language such as English, which has relatively impoverished morphology. It has been argued that morphological effects can be found in the absence of semantics for languages with more robust morphological systems (Plaut & Gonnerman 2000). The proposal within IM models is that morphological effects are found whenever there is sufficient evidence within the system for the morpheme to emerge. When languages have more robust morphological patterns, the language processor becomes more

sensitive to the patterns. In other words, if 1000 words all exhibit the same phonological pattern, those words can exhibit morphological behavior even if the meaning of all 1000 words is not constant. However, semantics is more crucial to the emergence of morphemes in a language like English, which consists of a morphological system with few fully productive morphological processes, many irregular processes, and morphological alternatives for the same semantic or grammatical function.

While most of the linguistic theories discussed above view these words as complex, they make very different hypotheses about how complexity is represented in the lexicon. Morpheme-based models and IA models assume that bound roots have independent lexical status; they have their own memory representation independent of the words in which they occur. Word-based models and some IP models would not allow for an independent lexical representation of the bound root. Only words, not morphemes, have permanent memory representations in word-based models, however a WFR can still serve to identify a word as consisting of an affix plus root. Thus, an investigation of bound root words may help to distinguish between some of the linguistic theories as well as suggest refinements to the theories to allow them to better account for the pattern of results produced by these words.

1.5.4 The role of form and meaning

For a long time, form and meaning were considered the crucial ingredients for a morpheme. However, bound roots posed a problem for this view. The claim that *receive* and *deceive* are not obviously related to each other suggests that an analysis of these words as complex is inconsistent with definitions of the morpheme that are dependent on meaning, such as Bybee's model, but a complex analysis of these words is not inconsistent with morphological theories that do not view meaning as a necessary component of a morpheme. However, the root is not the only source of meaning in prefixed bound root words; the prefix can also supply an

identifiable meaning component which can aid in identifying the word as polymorphemic. For example, the word *recede*, meaning "to move back or away", does seem to contain a meaning element consistent with the meaning of the prefix {re-}. The prefix is a source of meaning which has generally been overlooked in the psychological literature. I will contrast bound root and free stem words with both semantically transparent and opaque prefixes to see whether the transparency of the prefix influences the morphological information included in the lexical representations of prefixed words.

Just as bound root words often lack a semantically transparent root, they also often lack a phonologically constant root. Root morphemes like {-ceive} alternate with {-cept} when they undergo suffixation. If a morpheme is defined as a unit of form and meaning, then idiosyncratic phonological alternations of the sort in *receive ~ reception* could obscure the fact that these words share a root. In contrast, the fact that many words alternate in this same way, e.g., *deceive ~ deception, conceive ~ conception*, may aid in the identification of a common root morpheme in these words.   Thus, I will contrast prefixed words with alternating bound roots, like *receive*, to prefixed words with non- alternating bound roots, like *resist*, to see if the presence or absence of an alternation influences the representation of morphological structure.

1.6 Organization

In this chapter I have outlined the terminology needed to discuss morphological relations, presented some of the arguments for morphemes, and introduced some of the problems bound root words have posed for the traditional definition of the morpheme. I have also outlined some of the basic theories that have been proposed for morphological knowledge as well as word recognition models and set up the relevant issues to be investigated in this thesis.

In Chapter 2, I present an overview of the experimental evidence for morphological complexity. In Chapter 3, I present the methodology that was used in this dissertation and

present results from the first study using this task. Chapter 4 presents the first study investigating prefixed words. In Chapter 5, I present a follow-up study that investigates the role of phonological constancy. In Chapter 6, I present two experiments which investigate the influence of the semantic transparency of prefixes on the representation of complexity. Chapter 7 provides converging evidence from bound root and free stem words investigated with a different methodology. In Chapter 8, I suggest a model that captures the pattern of results and discuss their relationship to existing theories.

Chapter 2

EXPERIMENTAL BACKGROUND

Experimental evidence for whether or not lexical representations include any explicit morphological information comes from a variety of sources. Some of the evidence comes indirectly from researchers who are interested in the actual processes by which words are recognized. Other evidence comes from a variety of tasks aimed at uncovering evidence for morphological structure in the lexicon. Questions of morphological representations have been pursued for over 30 years; however there is still no clear answer as to whether morphological information is included in the memory representations of words. This is for a number of reasons. First, results have been quite inconsistent across methodologies. Second, the scope of what has been investigated has varied, in that the classification of a word as complex or simple across studies has been inconsistent. Thus even if one can say with certainty that morphological information is represented in lexical representations, identifying the set of words to which this statement applies is not straightforward. This statement could apply only to inflectional morphology, to free stem words, to suffixed words, or to semantically compositional words. Identifying where to draw the line between complex and simple words is as unclear as the more basic question of whether morphemes are needed at all.

In this chapter, I present some of the early evidence for morphological representations and then present some of the specific issues that have been addressed over the past 30 years. Specifically, I present studies that have investigated morphological effects and semantic overlap, morphological effects and phonological or orthographic overlap, affix priming, morpheme boundaries, and morphemes as an organizing principle within the lexicon.

2.1 Early Evidence for Morphemes

The first studies investigating the representations of morphologically complex words were morphological priming studies. In a priming study, the time needed for the recognition of a target word is affected by the prior presentation of a prime word, which is hypothesized to share some crucial feature with the target. Priming can be facilitative, meaning that the target word is recognized faster following the prime than in some baseline condition, or inhibitory, meaning that the target word is recognized more slowly following the prime than in some baseline condition. Features that have been shown to elicit priming effects are form similarity (both orthographic and phonological, e.g., Tanenhaus, Flanigan, & Seidenberg 1980) and meaning similarity (e.g., Swinney, Onifer, Prather, & Hirshkowitz 1979).

Priming methodologies are used in combination with a variety of tasks and presentations. Priming presentations can be visual, auditory, or a mix of both modalities. The amount of time that a prime is presented to a participant can vary, as can the interval between prime and target. The type of response to a target can also vary. Some tasks require a participant to read a target word aloud; this is called a **naming** task. Other tasks require a participant to decide whether a target is an actual word; this is called a **lexical decision** task.

Priming effects are usually understood in terms of activation models of word recognition. Put simply, the representations that are activated by the prime may have residual activation when the target is presented. If the prime and target share representational overlap, then the target may already be partly activated even before the input is perceived. Thus, its recognition time will be faster than if an unrelated prime had been presented. For current purposes, shared activation can be thought of as the result of either an overlap in lexical representations between prime and target or a lexical relationship between prime and target. For example, if two words sound alike, both may be activated by the same phonetic input; if two words mean the same thing, both will

be activated by the same conceptual node. If the phonetic string that corresponds to the word *cat* is detected, the beginning of that word will partially activate other words, such as *cab, calf,* and *cast,* in addition to fully activating the word *cat.* If the phonetic string that corresponds to *cab* is subsequently presented, the time needed to identify that word may be reduced if there is any residual activation left from the presentation of the word *cat.* The residual activation results in speeded recognition of a word or priming.

Some of the earliest evidence for morphological priming comes from Murrell and Morton (1974). They were interested in establishing whether the lexicon was morpheme or word-based.[10] They wanted to demonstrate a priming relationship between inflected forms of a stem. Using a long term (10 - 45 minutes) visual priming task, they found that an inflected word (e.g., *seen*) primed another inflected word with the same stem (e.g., *sees*) to the same degree as an identity prime, i.e., when the prime and target were identical, as when *sees* primes *sees.* Since the phonologically related control word (*e.g., seed*) did not prime *sees* at all, they concluded that the priming could not have been due to visual similarity. However, they could not rule out the possibility that this result was purely due to the meaning similarity between the two words.

Kempley and Morton (1982) followed up this question with an auditory priming experiment. They preceded a base target with its regularly inflected and irregularly inflected relatives, including suppletive forms. Since the irregular words share the same semantic relationship but do not share a morpheme (under the authors' assumptions that morphemes are defined as "a unique intersection of semantic and sensory characteristics" pg. 443), the absence of priming in the irregular condition would provide evidence that morphological effects are not reducible to semantic effects.

---

[10] For their purposes, morpheme-based means that words that are morphologically related will all be activated when the shared morpheme is presented. It does not necessarily mean that only morphemes are represented in the lexicon, which differs from what this term mean in the linguistics literature.

In this auditory study, primes were presented together in a (pre-experiment) training block in a clear, undistorted voice. Following the presentation of all the prime words, the target words were then presented in white noise such that their correct recognition was reduced to 40%. Participants were asked to write down each target word. Accuracy was much higher when the target word had previously been presented as an identity prime (e.g., *school* primed *school*) than when the prime was a phonologically similar but unrelated word (e.g., *school* did not prime *spool*). Accuracy was also high when a regularly inflected form of the target had been presented as a prime (e.g., *reflected* primed *reflecting*), but not when an irregularly inflected form of the target had been presented (e.g., *held* did not prime *holding*). Since the irregularly inflected primes had the same semantic relationships to the targets as the regularly inflected primes, the authors concluded that morphological effects were not reducible to semantic effects.

While these authors found no evidence of priming by irregularly related primes, this may have been due to the demands of the identification task, which requires participants to focus on the acoustic properties of the input in order to correctly identify the word in noise. Attention may have been focused entirely on the phonological characteristics of the items and semantic relatedness might have been ignored. Furthermore, it is interesting to note that these authors did not match the semantics or word forms across their regularly and irregularly inflected conditions. For example, their irregular condition contained eleven instances of comparative or superlative morphology, while the regular condition had only nine. The irregular condition contained 24 base forms (i.e., uninflected forms) out of 118 words while the regular condition contained 30 base forms. The target words in the two conditions were not the same either. For example, they did not contrast the ease of identifying a base like *hold* given the regularly inflected form *holds* to the irregularly inflected *held*. Each of these differences presents a minor confound in their study.

Kempley and Morton's study suggests that morphological priming is not restricted to the visual domain. It also suggests that morphological priming effects cannot be reduced to semantic priming effects. However, their investigation was restricted to inflectional morphology and the results may not extend to derivational morphology. The status of derivationally and inflectionally related words had been investigated in earlier research (Stanners, Neiser, Hernon & Hall 1979a). These authors used a relatively short term (10 intervening items) priming technique with a visual presentation and a lexical decision response to targets. They were interested in identifying whether derivationally and/or inflectionally related words shared a memory representation (comparable to a lexical representation) with their stems. Stanners et al. found that lexical decision times to an uninflected target word (e.g., *burn*) were equally as fast following an inflected form of the target (e.g., *burns*) as following an identity prime (e.g., *burn*). Irregularly inflected words (*hung ~ hang*) and derivationally related words (*selective ~ select*) produced less priming than the identity primes, although both conditions elicited significantly faster responses than in the unprimed baseline condition (e.g., *select* not preceded by anything). Stanners et al. concluded that regularly inflected words share a memory representation with their stem while irregularly inflected words and derivationally related words do not. However, the fact that they also succeeded in finding significant priming for both irregularly and derivationally related words at this short lag suggests that morphological priming is not restricted to regular inflectional relationships.

The contrast between Stanners et al.'s results (significant priming for both irregularly inflected and derivationally related words) and those reported by Kempley and Morton (no priming between irregularly inflected words) highlights the types of differences in results that are produced when the specific details of the task and presentation are changed. The former used a short term visual lexical decision task while the latter used a long term auditory word

identification task. Each task makes different demands on the attention of the participant. As discussed above, the fact that a word identification task with a long lag fails to find any influence of shared semantics between irregularly inflected forms of words is not surprising since the task focuses attention on the acoustic properties of words. The fact that Stanners et al. (1979a) did find that irregular inflectional relatives primed their bases is also not surprising since the lexical decision task requires participants to search their lexicon for an entry that corresponds to the target. It does not focus attention on a particular aspect of the word; however, it does force participants to make a judgment on lexicality, which requires them to set a threshold of acceptability which the identification task does not explicitly require.

Since the introduction of these early studies on morphological priming effects, many more studies have been conducted which demonstrate priming for both inflectionally and derivationally related words. However, the interpretations of these results have been debated. Some interpret these results as supporting the explicit representation of morphemes, which function as an organizing principle within the lexicon (Stanners et al. 1979a, Feldman 1991, Marslen-Wilson et al. 1994), while others have interpreted the results more as deriving from overlap between the form and meaning representations of words (Seidenberg 1987, Seidenberg & McClelland 1989, Rueckl et al. 1997, Plaut & Gonnerman 2000). Thus, depending on the interpretation of the results, conflicting models are supported or challenged by the results of these studies.

2.2 The Role of Semantic Overlap

Studies arguing against an independent level of morphological representation or organization within the lexicon look for evidence that morphological effects vary as a function of the degree of semantic and phonological overlap. Since morphological effects in their view have no status independent of the semantic and phonological relationships between words, they believe that the

magnitude of morphological effects should vary as a function of the degree of semantic or phonological overlap. As semantic or phonological overlap increases, the size of morphological effects should increase. As semantic and phonological overlap decrease, the size of morphological effects should decrease (Marslen-Wilson et al. 1994, Rueckl et al. 1997). Studies arguing for an independent representation of morphological complexity attempt to distinguish morphological effects from semantic and phonological effects. These studies have shown that morphological priming effects persist beyond either semantic or phonological priming effects (Bentin & Feldman 1990, Stolz & Besner 1998) and that morphological effects persist in the absence of one of these types of information (Stanners, Neiser, & Painton 1979b, Emmorey 1989, Napps 1989, Grainger, Colé & Segui 1991, Drews & Zwitserlood 1995, Stolz & Feldman 1995, Feldman & Soltano 1999).

Some studies have successfully distinguished morphological effects from semantic effects (Stanners et al. 1979b, Emmorey 1989, Frost, Bentin & Feldman 1990, Forster & Deutsch 1997). Emmorey (1989) is perhaps the only auditory priming study to find morphological priming effects in the absence of a semantic relationship between primes and targets. Using an auditory lexical decision task, she contrasted the priming effects found between bound root pairs (*e.g., submit ~ permit*) and unrelated morphologically simple words that rhyme (*e.g., saloon ~ balloon*). Finding more priming in the hypothesized morphologically related condition than in the phonological control condition, she concluded that morphological relations must be encoded independent of semantic and phonological representational overlap.

Stanners et al. (1979b) provide some of the earliest evidence of morphological priming in the absence of a strong semantic relationship. They used a visual priming lexical decision task to investigate the representations of prefixed words. They contrasted the amount of priming found for prefixed free stem words (*untrue*) and two types of prefixed bound root words: words for

which the root did not occur in any other forms (*rejuvenate*) and words whose root occurred in several forms (*progress*). They were interested in determining whether these words had unitary lexical representations or decomposed (i.e., represented as two separate morphemes) representations. Stanners et al. used two separate primes for each complex target: a root prime (e.g., *juvenate*) and a prefix prime (e.g., *remit* for *rejuvenate*). Their baseline comparison was the identity prime (e.g., *rejuvenate* for *rejuvenate*). The primes used for each target condition are presented in Table 1. The target words followed primes after 8 - 12 intervening items. The prefix prime was presented between the root prime and the target, also separated by several intervening items.

	Unique roots	Common roots	Free stems
Root prime	juvenate	regress	true
Prefix prime	remit	provide	unhappy
Identity prime	rejuvenate	progress	untrue
Target	rejuvenate	progress	untrue

Table 1: Primes and targets for Stanners et al.'s (1979b) three experiments.

Response times to target words after root and prefix primes were contrasted to the response times following an identity prime (*e.g., rejuvenate* priming *rejuvenate*). They argued that "any difference in the average latency between the critical target condition and the identity control condition would be due to a difference in the amount of priming provided by the stem and prefix primes as compared to the priming provided by the word itself" (Stanners et al. 1997b: 736) The same pattern of results was found for all three types of target words. Significantly more facilitation was found in the two-component prime condition compared to the unprimed condition but it was still significantly less than when compared to the identity prime condition.

Stanners et al. interpreted these results as evidence that there must be a unitary, unanalyzable, representation for complex words. If complex words had no lexical representation independent of their component morphemes, then the separate activation of each of the component parts in the two-prime conditions should have led to priming equivalent to that found in the identity condition. Despite the evidence for a unitary representation for complex words, Stanners et al. also found partial priming in their two- prime conditions.[11] They interpreted the partial priming as evidence for lexical relationships between the lexical representations of a prefixed word and its base morpheme. This study also provides evidence for priming between words sharing a bound root, which generally lack a strong semantic relationship. Furthermore, the magnitude of the priming effect for bound roots was equivalent to the magnitude of the priming effect reported for free stems.

The fact that they obtained equivalent results across their three conditions is somewhat surprising given that they used three very different root primes. The root prime for the unique root condition was the bound root by itself (e.g., *juvenate*), which is not an actual word. The root prime for the common root condition was a related derivational word, not the root alone, e.g., *regress,* not the root {-gress}. The root prime in the free stem condition was the free standing stem by itself (e.g., *true*), not a derived version of the stem (e.g., *truth*).

While the results of Stanners et al. (1979b) and Emmorey (1989) demonstrate a separation between semantics and putative morphological effects, many subsequent studies have failed to do so. The above results contrast with the results reported in the much-cited paper by Marslen-Wilson et al. (1994). These authors conducted a series of cross-modal (auditory primes and visual targets) lexical decision studies aimed at determining the relationship between morphological priming effects and phonological and semantic priming effects. They conducted

---

[11]Given the long interval of time between the presentation of the primes and the targets, they did not view this as the result of form priming.

their studies using a variety of morphologically complex words including prefixed words, suffixed words, free stems, and bound roots. To examine the role of phonological overlap, they contrasted derived words that have total phonological overlap with their bases (*friend ~ friendly*) to derived words that have partial phonological overlap with their bases (*elude ~ elusive*). They found no effect of reduced phonological overlap; the two phonological overlap conditions produced equal priming results.

To examine the role of semantic relatedness, they contrasted morphologically related words that were not semantically related (*successful ~ successor* or *depress ~ express*[12]) to morphologically related words that were semantically related (*punishment ~ punish* or *insincere ~ sincere*). Semantically related pairs produced significant priming while semantically unrelated pairs did not. Since significant morphological priming was not found in the absence of semantic relatedness, these results initially suggest that morphological effects can be reduced to semantic effects. However, one finding prevented them from abandoning morphological information in favor of semantic information, namely the fact that the presence of a semantic relation did not always result in priming either. In fact, Marslen-Wilson et al. (1994, as well as Marslen-Wilson, Zhou & Ford 1996) consistently failed to find priming when prime and target were both derived suffixed forms (e.g., *confession ~ confessor*), even when these words were semantically related. In contrast, they did find priming when the derived forms were prefixed (e.g., *unfasten ~ refasten*). To explain this result, Marlsen-Wilson et al. proposed a model of lexical representations, which includes inhibitory links between suffixed words derived from the same stem. Thus, while they found no direct evidence from priming that morphological effects were distinct from semantic priming effect, they still needed to include explicit morphological relationships between the lexical entries of suffixed words derived from the same stem to explain

---

[12]Note that the *depress ~ express* pair is the same type of word pair used by Emmorey in her auditory lexical decision study in which morphological priming effects were found.

the absence of priming between word pairs like *confession ~ confessor* or *governor ~ government*.

In summary, the investigation of the relationship between morphological and semantic priming effects has produced conflicting results. Some studies suggest that morphological and semantic priming effects are independent (Emmorey 1989, Stanners et al. 1979a, 1979b) while others suggest that morphological priming is dependent on semantic relatedness (Marlsen-Wilson et al. 1994).

2.3 The Role of Phonological Overlap

A similar inconsistency of results is found in the literature addressing the independence of morphological priming effects from phonological overlap. While some studies have failed to find any effect of reduced phonological overlap (Grainger et al. 1991, Marslen-Wilson et al. 1994, Drews & Zwitserlood 1995), others have succeeded in demonstrating that a decrease in phonological overlap results in a corresponding reduction in morphological priming effects (Emmorey 1989, Ruckel et al. 1997).

Many studies, including Marslen-Wilson et al (1994) reported above, have suggested that morphological effects are independent of phonological or orthographic overlap (Grainger et al. 1991, Drews & Zwitserlood 1995). However, other studies have found a correlation for the amount of phonological overlap between morphological relatives and the size of the priming effects. For example, Rueckl et al. (1997) used a primed fragment completion task to demonstrate a correlation between the degree of phonological overlap between morphologically related primes and targets and the strength of the priming effect. In this task, target words are presented with one character replaced with a hash mark (#). For each target word, there were at least four possible words that could have been created by replacing the hash mark with a single letter (e.g., the target word *mar#* could be *mark, Mary, Mars, mare)*. Participants read a long

study list of words (primes) and then (after a filler task), were presented with the masked targets. Participants decided which letter they thought was missing. Primes either included the intended target (*mark ~ mark*), a morphological relative (*marker ~ mark*), or an unrelated phonological control (*market ~ mark*).

Rueckl et al. compared the number of letter responses that matched the prime. They found the highest proportions of matches in the identity condition (*mark ~ mark*), followed by the morphologically related condition (*marker ~ mark*). The phonological control condition, (*market ~ mark*) produced no more matches than the unprimed condition. Following this study, Rueckl et al. conducted a second experiment which manipulated the degree of phonological overlap between prime and target. In this experiment they only used irregular, strong verbs which differed in a single segment (*make ~ made*) or more than one segment (*take ~ took*). They found a significant difference between prime/target pairs that differed in only one segment as compared with pairs that differed in more than one segment. They argued that this result shows a correlation between the strength of the morphological priming effect and the degree of phonological/orthographic overlap. They interpreted these results as supporting a connectionist, IM, model of morphological information, in which morphological information is not explicitly represented.

Studies which found a dissociation between phonological and orthographic overlap and morphological effects generally used naming and lexical decision tasks with very short lags between prime and target. The brief interval between prime and target may have reduced the influence of small phonological mismatches on the amount of priming. In contrast, Rueckl et al. used a fragment completion study with a very long lag between the primes, presented in a blocked study group, and targets. The very nature of the task, which requires participants to replace the hash mark with a letter to produce a word, may have directed participants' attention

to the orthographic form of the words and away from sources of information that may have otherwise been available from the prime presentations. In sum, investigations into the correlations between phonological overlap and morphological priming effects have produced the same sorts of conflicting results as the studies into the role of semantic relatedness.

2.4 The Representations of Affixes

Despite the competing interpretations of base priming effects as dependent or independent of phonological and semantic overlap, their existence has been firmly established. However, there has not been the same supply of evidence for an affix priming effect. In fact, there has been very little (if any) evidence to suggest that words that only share an affix are lexically related.

In addition to finding evidence for morphological priming of roots independent of a semantic relation, Emmorey (1989) also conducted auditory lexical decision experiments comparing the priming effects for words which shared an inflectional or derivational affix (*winking ~ poking,* or *blackness ~ shortness*) compared to morphologically simple words that just happen to share a final string (*tango ~ cargo*). She found no evidence of any morphemic priming. Rather, words that shared a full final syllable (e.g., *tango ~ cargo, winking ~ poking,* or *blackness ~ shortness*) produced significant priming, regardless of morphemic status or type of affix, while words that shared a morpheme that was not co- extensive with a syllable (e.g., *smiling ~ breaking*) did not. Furthermore, the magnitude of the priming in the phonologically-related condition was equivalent to the magnitude of priming in the suffix condition; thus no additional priming was found for sharing an affix.

Other researchers have also failed to find significant priming for words that share affixes. A particularly interesting example of this comes from Frost, Forster and Deutsch (1997). They compared the amount of priming found for consonantal root patterns versus affix-like vowel patterns in Hebrew using primed lexical decision and naming tasks. In their studies, primes

were briefly visually presented (43 ms) such that participants had limited conscious knowledge of the prime. Targets were presented immediately after the prime. In both lexical decision and naming tasks, these authors failed to find any priming effect for words which shared a vowel pattern while they did find facilitated responses to targets when they were preceded by the three letters that formed the consonantal root.

What explanation might there be for root morphemes to exhibit priming effects while affixes do not? Affixes are similar to root morphemes in that they maintain a consistent form across words and have consistent semantics. Furthermore, since many affixes are highly productive, they can be found on hundreds of words, much more than the average root morpheme. Possibly, the absence of affix priming effects is due to the amount of form and meaning contributed by the affix relative to the base. In most cases, the base of the word provides its core meaning. In comparison, the semantic contribution of an affix is slight. Affixes in English also tend to contribute less of the phonological material than the base of the word. The average affix in English is at most 2 or 3 characters in length. Word roots, on the other hand, are often 4 to 5 characters long, and can be longer. In terms of facilitating the recognition of a subsequent word, affixes simply may not provide enough semantic and phonological information to sufficiently prime a word that shares the affix. In an affix priming study, primes and targets will necessarily have different bases (e.g., *reheat* and *reread*). The base provides the core of the semantic and phonological information. Any facilitation due to residual activation of an affix might be overpowered by the competition between the bases. Thus, failure to find affix priming effects does not necessarily indicate that affixes do not have an independent representation, that words which share an affix are not lexically related, that the prior presentation of an affix does not leave residual activation on all words that shares the affix, or that the affix might not also affect the processing and representation of the word. However, the frequent failure to find priming

effects for affixes suggests that an investigation of their representation with an experimental task not dependent on priming might be more fruitful.

Several non-priming methodologies have been described in the literature. Naming and lexical decision tasks both have non-primed counterparts, for example. They have both been used extensively to investigate the representation and processing of morphologically complex words. For example, Wurm (1997, 2000) investigated the contribution of affix characteristics with several methodologies, including an unprimed lexical decision task. In his experiments, Wurm contrasted the lexical decision time to nonce words composed of real or fake bound roots, free stems, and prefixes. Example stimuli are provided in Table 2. For both the real root and the fake root conditions, words with real prefixes elicited longer judgment times than words composed of fake prefixes. This finding is consistent with prior lexical decision tasks which compared nonce words composed of actual morphemes to nonce words composed of non-morphemic strings, as will be discussed in section 2.5 (e.g., Taft, Hambly & Kinoshita 1986, Caramazza, Laudanna, & Romani 1988, Allen & Badecker 1999).

	Real Prefix		Nonce-Prefix	
	real base	nonce-base	real base	nonce-base
Free stems	adlay	adloo	aflay	afloo
Bound roots	adlect	admect	aklect	akmect

Table 2: Sample stimuli from Wurm (2000).

In addition to investigating the influence that real morphemes have on the lexical decision times for nonce words, Wurm also investigated specific characteristics of the prefixes in his stimuli. He was interested in whether characteristics of the prefix could affect lexical decision times to nonce words composed of real morphemes. He developed a number of measurements, such as Prefix likelihood, Prefix frequency, and Root frequency. Prefix likelihood was defined as

the ratio of the "summed frequency of the truly prefixed words beginning with a given phonetic string" over the "summed frequency of all words beginning with that string in which removal of the string leaves a pronounceable syllable or syllables" (Wurm 2000:258). For example, the word *reach* would not be included in the calculation of Prefix likelihood of the prefix {re-} because when the initial string, [re-], is removed, the remainder of the word, [_], is not a full syllable. It should be noted that only fully compositional free stem words are included in Wurm's set of 'truly prefixed words'. Prefix frequency was calculated by summing the counts for each word beginning with a given prefix.

In addition to the faster response times for nonce words composed of real prefixes but fake roots, Wurm also found that both root and prefix frequency was inversely related to judgment times. Nonce words with high frequency prefixes had longer response times than those with low frequency prefixes. Nonce words with high prefix likelihoods also had longer response times. Furthermore, the cost in processing time was more severe when a prefix with a high prefix likelihood value co-occurred with a free stem than a bound stem. For example, the prefix {trans-} has a relatively high prefix likelihood value of .092. Thus, the response time for a free stem nonce word with this prefix (e.g., *transcut)* was longer than the response time for a bound root word with this prefix (e.g., *transflict*). Furthermore, both of these words would have longer response times than the corresponding root with a low likelihood prefix such as {de-}.

These results suggest that characteristics of a prefix, such as frequency and prefix likelihood, can affect the way in which a word is processed and/or represented. Wurm argues that these affix traits interact to identify a word as a good candidate for lexical decomposition.

2.5 Evidence for Morpheme Boundaries

While the majority of the studies reported above used a priming methodology, there are several other tasks available to the researcher for the investigation of morphological

representations in the lexicon. Since priming is argued to be the result of shared representational overlap or links between lexical representations, it cannot discriminate between models which propose that words have unitary representations with links between morphological relatives and models which propose that words have complex representations with internal morphological structure explicitly represented. Nor can priming tasks discriminate between IM and EM models.

The nonce word lexical decision task used by Wurm (2000) has also been used to argue for the inclusion of morphological structure in lexical representations. Lexical decision times are compared for nonce words consisting of either real morphemes in allowable combinations (e.g., *happyment), non-morphemic strings that differ from morphemes by a single segment (e.g., *happuniss), or one real morpheme and a nonsense string that differs from a real morpheme by a single segment (e.g., *happyniss or *happuness). Research using this task has repeatedly found that nonce words constructed from real morphemes take longer to reject than nonce words constructed from nonsense strings (Taft et al. 1986, Caramazza et al. 1988, Allen & Badecker 1999, Wurm 2000). Likewise, judgement times to nonce words partially composed of real morphemes are longer than judgement times to nonsense strings but not as long as judgements to nonce words constructed of all real morphemes. Using similar materials, researchers have found the opposite result using a nonce word naming task (Laudanna, Cermele, & Caramazza 1997). Namely, nonce words consisting of real morphemes are named faster than non- morphemic strings. The fact that lexical decision and naming lead to opposite results is consistent with a model of the lexicon in which words are represented with internal morphological structure. Real morphemes make a nonce word harder to distinguish from real words but easier to pronounce, since they do not need to be sounded out from the orthography. If words were represented in an undecomposed fashion, having nonce words that were constructed from real morphemes should

not affect either the naming or lexical decision times. These words would be treated just as any unstructured nonsense word.

Another question investigated by Libben (1993) with this sort of nonce word task is whether relationships between the morphemes of complex words are maintained in their representations, as proposed by certain linguistic theories (Lieber 1980, 1992, Selkirk 1982). Specifically, Libben (1993) investigated this question by combining nonce roots {ponk} with real affixes. He joined nonce roots with suffixes and prefixes in allowable and unallowable combinations, as determined by the selectional restrictions of the affixes. He contrasted the naming times for nonce words with disallowed morpheme combinations (*reponkity*) to nonce words with allowable combinations (*reponkize* and *reponkable*). Note that in the disallowed forms, the suffix {-ity} attaches to adjectives and produces nouns. However, the prefix {re-} only attaches to verbs. Thus, the combination of re-X-ity is never allowed, regardless of the meaning of the stem. Libben found that disallowed combinations of morphemes (*reponkity*) took longer to name than allowable combinations. This result provides strong evidence that morphological structure plays a role in word recognition, and that there must be some morphological structure included in lexical representation or calculated during recognition.

While the results of these nonce word studies are fairly consistent and compelling, the primary criticism leveled against them is that one cannot necessarily generalize results based on nonce words to real words. It is possible that nonce words are processed very differently than real words and therefore the results are not directly relevant or informative to the question of how real words are recognized and represented. In fact, some of the authors who conducted this research propose different recognition procedures for nonce words than for real words (Caramazza, Miceli, Silverri & Laudanna 1985, Caramazza et al. 1988).

Another task that has been used to argue that morphological structure is encoded in lexical representations and used by the language processor is the segment shifting task (Feldman & Fowler 1987). This task, in which participants remove a string from one word and attach it onto another, has been used to demonstrate that suffixes in English and Serbo-Croatian are more separable from the whole word than pseudosuffixes, non- morphemic elements that are string identical to a matched morpheme (Feldman 1991, Stolz & Feldman 1994, Feldman et al. 1995). For example, Feldman and her colleagues found that the {-en} in *harden* is more separable from the word than the matched phonetic string in the word *garden,* which is not a suffix. Again, these results support a view of the lexicon in which words are represented with internal morphological structure.

Similar to the segment shifting task is a word recognition task with a luminance manipulation (Beauvillain & Segui 1992, Beauvillain 1994). In these experiments, conducted in French, the time to recognize a word is contrasted across three luminance conditions: all high luminance (**reflux**), all low luminance (reflux), and partial high luminance (re**flux**). They also contrasted two types of words in the partial high luminance condition: complex word where the root is highlighted (re**flux**) and a morphologically simple word with a pseudoprefix in low luminance and pseudoroot in high luminance (re**flet**). Beauvillain argued that high luminance characters are perceived faster (by the retina) and therefore should speed recognition if the unit that is in high luminance corresponds to a lexical unit. These authors found that the times required to recognize the words were faster in the partial luminance condition than in the low luminance condition only when the highlighted portion of the word corresponded to a real root. When the highlighted portion was a pseudoroot, no benefit of high luminance was found. Beauvillain interpreted these results as evidence that the root is the basic unit of access for words

in reading. This result also suggests that morphological structure is encoded for complex words in the lexicon.

2.6 Morphemes as an Organizing Principle in the Lexicon

The last type of evidence for the inclusion of morphological information in the lexicon comes from studies of what is termed the 'morphological family effect'. In essence, a morphological family consists of all the words in a language that are derived from the same root. The morphological family effect finds that words with large families are responded to more quickly than words with smaller families (Meunier & Segui 1999, Bertram, Schreuder & Baayen 2000a, Bertram, Baayen, & Schreuder 2000b). The studies investigating morphological family are usually unprimed naming or lexical decision tasks. Two conditions are compared, a large family condition and a small family condition. Family size is defined or calculated very differently across studies. Some use a token count, while others use a type count. Also, some include only semantically related words. However, despite the differences in how family size is calculated, the result is quite consistent; large families speed recognition of their members. To account for this finding, models of the lexicon have been proposed in which all words sharing a base are lexically connected. This result supports the notion that lexical entries are organized into morphological paradigms in the lexicon. It should be noted that analogous studies have been conducted using a phonological family size with the opposite effect. Words that have a large phonological family size (or a dense phonological neighborhood) elicit slower response times than words with smaller phonological families (Goldinger, Luce & Pisoni 1989). Thus, the morphological family size effect cannot be reduced to the phonological similarity that the members of a family share. There are no analogous studies investigating semantic features, as far as I know.

To summarize the findings presented above, there is considerable evidence in the literature that words are represented with internal morphological structure. The results from non-priming tasks have provided clear evidence in support of this claim. The morpheme is a unit which aids recognition, as was demonstrated by the nonce word studies of Caramazza and colleagues, Wurm's studies on prefixed words, Feldman's segment shifting studies, and Beauvillain and colleagues experiments using the luminance manipulations. Furthermore, evidence of a base priming effect have also been consistently reported. The problem is that experiments trying to distinguish morphological effects from phonological overlap and high semantic relatedness effects have produced mixed results. Finally, the morphological family size effect also supports a network of morphological relations, at least for semantically related family members. However, bound roots and affixes have received relatively little attention in the literature and there are many issues related to them that are still unaddressed.

Chapter 3

EXPERIMENTAL METHODOLOGY

The prior chapter discussed at length the various sources of evidence for the inclusion of morphological information in the mental lexicon. However, criteria for distinguishing complex words from simple words are inconsistent across experimental investigations and the specific models of how morphological information is represented differ across studies as well. To provide additional insight into how morphological information is represented in the mental lexicon, I will be contrasting the representations of morphologically complex words derived from semantically transparent free stems, such as the examples in (30) and semantically opaque bound roots, such as the examples in (31). Bound root and free stem words will also be compared to morphologically simple control words consisting of pseudoprefixes such as those in (32). I will present results from five segment shifting studies and one speech error elicitation study, all of which suggest that free stem words are represented differently from morphologically simple words. The results of these studies will also show that when a bound root word has a semantically transparent prefix, it has a representation that is distinct from morphologically simple words.

(30) reheat, cooperate, prejudge
(31) receive, cohesive, preside
(32) religion, coconut, precinct

The investigation of prefixed free stem words is interesting because many word recognition theories have argued that the processing of prefixed and suffixed words are different (e.g., Taft & Forster 1975). Specifically, it has been argued that word recognition is root-driven. To identify a word, one must access the lexical representation for the root before the prefixed word representation can be found. Since prefixes precede and camouflage the root, they can interfere with recognition in a way that suffixes may not. Bound root words have been underinvestigated relative to free stem words. In fact, many studies of morphological processing and representation ignore the

representation of bound root words. For example, many studies classify bound root words together with morphologically simple words (e.g., Wurm 1997). Moreover, as discussed in Chapter 2, the few studies that have investigated bound root words have found conflicting or equivocal patterns of results (e.g., Emmorey 1989, Marslen-Wilson et al. 1994). Furthermore, because bound root words are rarely semantically compositional, EM and IM models make different predictions as to how they should pattern in behavioral tasks that are sensitive to morphological information.

Most models of word recognition expect free stem words, such as those in (30), to be primed by their bases, although the explanation for the facilitation varies across theories. EM models explain morphological facilitation either as the result of explicit morphological representations that are shared by two words or as the result of associative links between morphologically related words. IM models explain morphological facilitation as the result of shared phonological and semantic activation patterns between the derived word and its stem. Despite these different explanations based on different representational assumptions, both models predict facilitation.

In contrast, EM and IM models make different predictions for bound root words, such as those in (31). Specifically, IM models that conceive of morphological information as emerging from shared phonological and semantic information do not expect words that share a bound root to prime each other unless they also share a strong semantic relationship. For example, although the words *receive* and *conceive* share the bound root {-ceive}, they are not obviously semantically related. Thus, IM models would not expect *receive* to prime *conceive* beyond the facilitation attributable to their phonological similarity. Rather, words with bound roots should behave as though they were morphologically simple words which just happen to have the same ending, like the monomorphemic words *saloon* and *balloon*. The alternative view, espoused by most linguists and consistent with EM models, is that combinations of abstract linguistic patterns such as formal alternations, subtle semantic relationships, string overlaps, etc. provide sufficient evidence for a speaker to encode a

morphological relationship between two words. Since *receive* and *conceive* both alternate in the same idiosyncratic way, namely {-ceive} alternates with {-cept}, and share a final string, they are morphologically related and thus predicted to exhibit morphological priming effects. Although many linguists consider both free stem and bound root words to be morphologically complex, the specific mechanism used to encode the morphological complexities for the two types of complex words may not be equivalent. However, they are both predicted to have representations that are distinct from the representations of simple words.

3.1 Criteria for Morphological Classifications

I will be comparing the behavior of free stem, bound root and simple words in a behavioral task that is sensitive to morphological structure in lexical representations. I hope to uncover evidence for the explicit representation of morphological information in the mental lexicon in both free stem and bound root words. But before any experiment can examine differences in representations and processing of these three classes of words, I must first make clear how I define these categories. In English, the distinction between bound root and free stem is no more obvious than the distinction between complex and simple.

3.1.1 Distinguishing complex words from simple words.

To distinguish between morphologically complex words and morphologically simple words, I use a combination of criteria. For my purposes, a simple word should be consistent with the following criteria: (a) it should not have an embedded word that is semantically related to the meaning of the whole word, (b) it should not have a substring that occurs with other prefixes, and (c) it should not have a meaning component that is consistent with the meaning of the prefix, or if it does, it should not participate in many derivational processes. The example in (33) exemplifies a violations of criterion (a).

(33) a.　　install
　　b.　　delight, relate, person

The word in (33a) looks as though it were composed of the prefix {in-} and the stem {stall}, but it is unclear whether it is actually derived from *stall* or whether this apparent relationship is more of a folk etymology. The meaning of the word is not compositional nor does the root occur with other prefixes (there is no word *prestall, *restall, or *destall). But the meaning of the embedded word *stall* could be viewed as contributing to the meaning of the whole word. One definition for *install* is 'to place in an office or dignity by seating in a stall or official seat'. Since *stall* plays some role in the meaning of this word, albeit slight, I did not consider this word simple. However, I also did not use *install* or similar words as instances of a complex word in any experiment. Contrast the example in (33a) to the words in (33b) which also have embedded words, but they contribute nothing to the semantics of the whole word. *Delight* is not semantically related to *light* and *relate* has nothing to do with being *late*.

Additional violations of my criteria are illustrated by examples (34)-(36).

(34) discrete ~ concrete
(35) retrieve
(36) premier

Example (34) provides a pair of words which individually meet criteria (a) and (c) in that they have neither an embedded word nor a meaning component that is consistent with either prefix, but together they violate criterion (b). The fact that both words have the same substring and a prefix-like initial syllable is sufficient to eliminate them from the pool of morphologically simple words from which items will be selected. Example (35) violates criterion (c) because the meaning contributed by the prefix, "back" or "again" is strongly present in the meaning of the whole word, "to get back", and the word participates in other derivational processes (*retrieve ~ retrieval ~ retrievable*). Example (36) also violates the first clause of criterion (c). *Premier*, defined as "first in position, rank or importance", has a meaning consistent with that prefix, {pre-}, which means

56

"before". However, *premier* does not undergo any other morphological operations: it has no

morphological neighbors, nothing is derived from it, and the 'root' {-mier} does not occur in any

other words. Thus, example (36) does not violate both clauses of criterion (c) and therefore would

be classified as simple.

Generally, words that are complex, i.e., have already undergone a morphological operation, are

open to other morphological operations, if other derivational processes are appropriate. (Note that a

complex word like *unconstitutionality* may have exhausted the derivational processes that it can

undergo.) In contrast, morphologically simple words often do not undergo any morphological

derivations, such as the words in (37). While it is certainly not true that morphologically simple

words in general do not undergo derivations, it is true that more simple words than bi-morphemic

words forego morphological derivation.

(37) a.   premier
     b.   Indian
     c.   contrast
     d.   person

All the words in (37) are classified as simple according to the criteria outlined above.

If a word had one or more of the characteristics outlined above, I either classified it as complex

or did not use it at all. While these criteria allow for consistent classification of words as either

complex or simple, they still leave some unresolvably fuzzy edges, as I will demonstrate in the next

section.

3.1.2 Distinguishing bound root words from free stem words.

A complex word that is composed of a prefix plus an embedded word that is semantically related

to the whole word is classified as a free stem word. If the word does not have an embedded word

and the word does not pass the above criteria for morphologically simplicity, then it is classified as a

bound root word. Problems arise for the distinction between bound root and free stem words when

words contain an embedded word that is **not** clearly related to the meaning of the whole word. If the embedded word is unrelated to the meaning of the whole word, then I usually classified it as a bound root word. This is especially true if the substring co-occurs with many prefixes, since bound roots tend to co-occur with a wider variety of prefixes than free stems (although the set of prefixes a bound root co-occurs with is essentially fixed while a free stem can participate more readily in neologisms), as illustrated in examples (38) - (39).

(38) a.   advent ~ convent ~ prevent ~ invent ~ circumvent ~ event
     b.   refuse ~ confuse ~ profuse ~ infuse ~ diffuse ~ suffuse
(39) a.   vent ~ ?re-vent ~ ventilate ~ unventilated
     b.   fuse ~ defuse ~ unfused ~ re-fused

Example (38) presents words composed of bound roots which are homophonous with free stems. In each case, the complex words are unrelated to the meaning of the free stem. For example, the free stem {vent} is defined in several ways, most incorporating the notion of expulsion or outlet of smoke, wind, emotions, etc. The words in (38a) do not express this meaning while the words in (39a) do. Furthermore, note how many more prefixes co- occur with the bound root {-vent} in (38a) compared to the free stem {vent} in (39a). Likewise, the words in (38b), derived from the bound root {-fuse} have little semantic overlap with the meaning of the free stem {fuse}, in (39b), which has few morphological derivatives. In examples like these, the distinction between free stem words and bound root words is based on the degree of semantic overlap with the base and on distributional evidence to suggest two separate morphemes, such as the contrast presented above for {vent} and {-vent}.

Even trickier is when an embedded word occurs with many prefixes, suggesting that it is a bound root, but in some cases it is semantically related to the whole word, suggesting that it may be a free stem.

(40) a.   depress ~ repress ~ oppress ~ express ~ compress
     b.   deport ~ import ~ report ~ comport ~ export
     c.   propose ~ impose ~ depose ~ oppose

In examples (40), the meanings of the embedded words are more related to the meanings of the whole words than the embedded words from examples (38) and (39). For example, under certain interpretations the words in (41a) could be seen as compositional and transparently related to the free stem {press}. *Depress* can mean "press down", *repress* has a sense of "press back", *compress* could mean "with pressure", and *express* could mean "press out". However, although these meanings are present, they do not capture some of the synchronic meanings of the words, many denoting psychological states. Furthermore, the free stem {press} has different morphosyntactic properties than the prefixed words, as the comparison between (41a) and (41b-c) shows.

(41) a.　press $_v$ ~ pressure $_n$
　　 b.　depress $_v$ ~ depression $_n$, *depressure $_n$
　　 c.　repress $_v$ ~ repression $_n$, * repressure $_n$

Words of the type in (41) are difficult to classify. However, the criteria of comparing semantic relatedness to the stem and comparing distributional and morphosyntactic patterns were applied in these circumstances and generally resulted in these words being classified as bound root words, not free stem words.

　Another complication is posed by words whose substrings are free standing words that are related to the whole word, but which are of very low frequency or are archaisms, as in (42).

(42) a.　subscribe ~ scribe
　　 b.　decline ~ cline

The words in (42) do have free standing roots which seem to be marginally related to the meanings of the whole words. However, the complication is that the root is either archaic and no longer used in the language or the meaning of the word has drifted significantly such that the relationship has diverged, making the words in (42) essentially like the words in (40). For instance, the word *scribe* no longer expresses the general sense of 'writing' but rather describes a very specific type of individual who copies manuscripts. Thus, its relationship to *subscribe* is very distant. Likewise, whether people are even aware of the word *cline* is unclear. In fact, some dictionaries only list it as a

bound root. Usually, I tried to avoid the use of these words in experiments. However, the set of prefixed words derived from bound roots is a finite set and therefore I was often faced with these types of judgments. Using such words was sometimes necessary. When they were used, they were always included in the bound root category.

Now that I have outlined my criteria for distinguishing between the three classes of word that will be investigated in this dissertation, I now turn to a discussion of the task that I will be using for the majority of my experiments.

3.2 The Segment Shifting Task

The majority of the experiments in this dissertation use the segment shifting task (Feldman & Fowler 1987). This task exploits the fact that the same orthographic and phonetic sequence can be morphemic in one word but non-morphemic in another. For example, consider the words in (43). Both of these words have the same word ending. However, the ending {-en} serves as a suffix only in (43a); in (43b) the same orthographic and phonetic string, -en, serves no morphemic function. The fact that these words have the same ending is purely coincidental.

(43) a.   harden
     b.   garden

The segment shifting task interrogates speakers' lexical representations by having them create words from the parts of other words. The task was inspired by patterns of laboratory-induced and naturally occurring slips-of-the-tongue that involve morphological components.   The logic underlying this task depends on the relative independence of morphemic elements in word representations. Feldman (1991) reasoned that if word components can be separated and exchanged, or omitted, in the course of word production, then perhaps the degree of separability of the components will be reflected in response times in an on-line task.

The general procedure is as follows: Participants are presented with a morphologically complex or simple **source** word. For example, COOPERATE is considered a complex source word since it is derived from the root OPERATE. In contrast, COCONUT is a simple source word since the CO-initial string is a pseudoprefix and -CONUT is not a base morpheme.[13] Source words are visually presented on a computer screen for 750 ms, at which point the (pseudo)affix is highlighted simultaneous with the presentation of a **target** word (AUTHOR). Participants are instructed to shift the highlighted portion of the source word onto the target word and pronounce a **response** word (COAUTHOR). The highlighted source word and target word are displayed for 1500 ms at which time the screen is blanked. The dependent variable in these studies is the time required to pronounce the response word. The shifting time comparison across conditions is always measured for the production of the same response word. What differs across conditions is the morphological structure of the source word. In other words, the time required to produce the word COAUTHOR when it is preceded by the complex source word COOPERATE is compared to the time required to produce the same word when it is preceded by the simple source word, COCONUT. Feldman reasoned that "if the morphological structure of source words is analyzed and decomposed in the course of segmenting the designated segment and building up the utterance to be named, then morphological effects are anticipated" (Feldman et al. 1995:949). Thus, the predicted pattern of results is that morphologically complex words should facilitate shifting relative to simple words since a real morpheme is a coherent unit that has some degree of independence from the base while a non-morphemic string has neither of these characteristics. Figure 2 schematically illustrates the time course for this procedure.

---

[13] The fact that *coconut* could be analyzed as consisting of the morphemes {coco-} and {nut} is irrelevant to this comparison. The crucial observation is that [co-] does not constitute an actual morpheme and [conut] is not a word base.

+	COOPERATE	<u>CO</u>OPERATE AUTHOR	COAUTHOR
Fixation point 200 ms	Source word 750 ms	Prefix highlighted target word presented for 1500 ms Clock starts	time to onset of vocalization measured
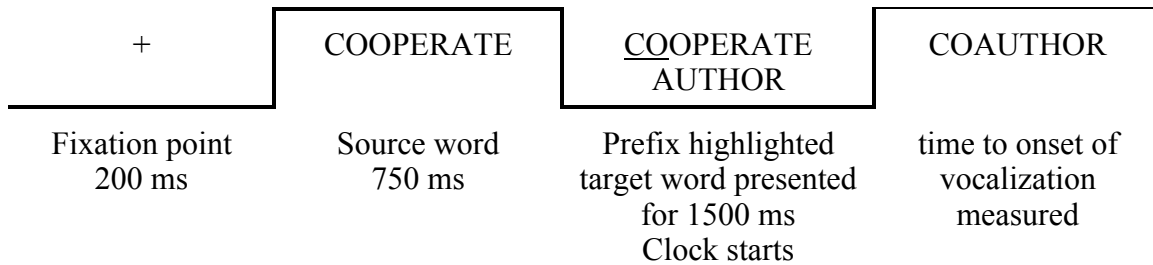
Figure 2: Procedure for segment shifting task designed by Feldman and colleagues.

Feldman and her colleagues have investigated morphological complexity in a number of different languages. They have shown a facilitatory effect of complexity for English inflectional and derivational suffixes (Feldman, Frost & Pnini 1995, Stolz & Feldman 1995), Serbo-Croatian inflectional, but not derivational, suffixes (Feldman 1991, 1994), and Hebrew inflectional vowel patterns (Feldman et al. 1995). However, to my knowledge this task has never been used to investigate prefixed words or a difference between free stems and bound roots.

3.2.1 The locus of the effect

Feldman and her colleagues interpret the results from their segment shifting studies as evidence that morphological structure must be encoded in the lexical representations of words. They argue against a processing explanation of the results on the grounds that source words are presented for a relatively long time before any operations must be performed by the participant. They argue that the long presentation duration for the source word "works against a prelexical account of morphological processing by giving virtually unconstrained processing time as lexical access is generally available before 750 ms has transpired." (Feldman et al. 1995:951, however see Balota & Chumbley (1985) for arguments against a view that recognition has *finished* within any fixed period of time). Furthermore, they maintain that factors which normally affect word recognition time are orthogonal to the segment shifting task. For example, they have found that segment shifting differences are generally immune to token frequency differences between conditions, which is a well-known correlate to word recognition times. This also suggests that the differences in shifting times are not

due to prelexical processing associated with the time required to recognize the source words but rather with the differences in the representations of the words. However, I should note that the segment shifting studies presented in the current thesis were affected by gross differences in token frequencies of the source word. Thus, it is not clear the extent to which this argument holds true, at least for prefixed words. Furthermore, certain models of lexical processing and representation propose that frequency differences impact the way in which lexical and sublexical units are actually represented (e.g., Grossberg, Boardman & Cohen 1997); thus the effects of frequency differences are not limited to prelexical processes, as Feldman suggests.

The segment shifting procedure requires the participant to perform two separate tasks. First, she must separate the highlighted string from the source word. Second, she must attach the string onto the target word. Thus, the locus of the segment shifting effect could come from either or both of these components. The difference in shifting times between morphemic and non-morphemic units could be attributed to the separability of an affix from its base or it could be attributed to the ease with which the shifted material can be added to the target word. The ease of attachment could be affected by a mismatch in the role the shifted material plays in the source word relative to the role it plays in the response word. Alternatively, the differences in shifting time could be attributed to a combination of these factors.

Feldman and her colleagues argue that the pattern of results can be attributed only to the separability of the affix from the whole word representation for the source word and not to the ease of attaching it to the target word. They suggest that separability is determined partly by the relative independence of affixes and stems and partly by the coherence of a real affix compared to the coherence of an arbitrary phonological string. Coherence of an affix refers to the degree to which a phonological string behaves like a unit. For example, in a word like *religion,* the string [re] behaves more like a unit than the string [rel] because the former corresponds to a syllable in the word while

the latter does not. Likewise, the string [re] in the word *reheat* should behave more like a unit than

the [re] of *religion* if it corresponds to a morpheme in the lexical representation of the word.

Evidence that the locus of the segment shifting effect is in the segmentation of the source word

rather than in the ease of attaching the segmented string to the target word has a number of sources.

First, Feldman (1991, 1994) investigated the influence of affix mismatches between source and

response words. She found no difference in shifting times for an affix when its function on a source

word was consistent (same syntactic category and gender) or inconsistent with its function on a

response word. For example, the Serbo-Croatian suffix {-i} is homophonous, functioning both as

third person singular and as the nominative plural. Feldman contrasted the time to produce the

response word, given in (44c), when the suffix had its morphemic source from a **consistent source**

word in which the string served the same morphosyntactic function, (44b), or an **inconsistent**

source word in which the string served a different mophosyntactic function (44a). If the

differences in shifting time were due to processes involved with attaching the affix to the target

word, then the shifting times in the inconsistent affix condition should have been longer than in the

consistent affix condition since the functional mismatch would need to be resolved. Instead, she

found no difference in response times between the two source word conditions. Whether the shifted

string served the same function in the source and response words or a different function did <u>not</u>

influence the shifting time. The absence of a shifting time difference between these two conditions

suggests that the locus of the effect is not in the degree of consistency between the affix in the

source word and the affix in the response word but rather in ease of separating the unit to be shifted

from the source word.

(44) a.   CEVI   'pipes'   *Consistent source word*
          CEV-I   pipe-pl
     b.   CEDI   'He wrings'   *Inconsistent source word*
          CED-I   wring-3sg
     c.   RADI   'He works'   *Response word*
          RAD-I   work-3sg

This result also suggests that the effect cannot be due to semantic priming between the affix shared by the source and response words. The consistent source words had more semantic overlap with the response word than the inconsistent source words. If the speeded shifting times exhibited by complex conditions were actually due to semantic priming, then these two conditions should have exhibited a difference, which they did not. Thus, this experiment provides initial evidence that the segment shifting task is sensitive to differences in the separability of the shifted string from the source word and not to the compatibility of the affix on source and response words. However, this finding was restricted to inflectional suffixes, which are more functional than conceptual.

Additional evidence that the segment shifting task is sensitive to differences in separability comes from the fact that shifting time differences between complex and simple source words are maintained even when target (and response) words are not real words. Feldman et al. (1995) hypothesized that if their findings were due to similarities between source and response words then the effect should be lost when the targets were not real words. To rule out this possible interpretation, they successfully replicated their findings using pseudowords as targets. They concluded that the difference in shifting times could not be attributed to similarities between source and response words. Also, since the response words were nonsense words lacking semantic and morphosyntactic features, this finding also argues against a semantic similarity explanation for the differences in shifting times.

While there is no evidence to suggest that source-target word compatibilities influence affix shifting behavior in this task, Feldman et al. (1995) did find that shifting times were correlated with "morphological reliability". The "morphological reliability" measure is similar to Wurm's Prefix likelihood measure, described in section 2.4. Morphological reliability is the ratio of words which end in a morpheme-like string to those in which the string actually is morphemic. They found that "the magnitude of the difference in shifting latencies was correlated with a crude measure of

65

morphological reliability for each affix" (Feldman et al. 1995:952). The fact that shifting times

were correlated with morphological reliability is also consistent with the hypothesis that segment

shifting is sensitive to the separability of the affix from the source word. The greater the probability

that a given string is an affix the more it will cohere as a unit and the more evidence to support its

independent representation.

The following quote summarizes Feldman's view on the task:

> The segment shifting task demands segmentation of source words. A match between the
> (experimentally induced) components of the source word and the components of that word
> specified lexically appears to facilitate performance. For morphologically simple (and
> monogamous) source words, segmentation and affixation of the final sequence of letters is
> difficult because it is linguistically arbitrary. For morphologically complex (and polygamous)
> source words, by contrast, segmentation and affixation is relatively easy because it is
> principled and may depend on units made more salient by their tendency to combine to form
> many different words (1995:958).

The segment shifting task was intended to distinguish between morphemic and non- morphemic

units, not to identify whether two words shared representational overlaps. As a result, Feldman and

her colleagues interpret segment shifting results in terms of an EM model of morphological

representation. However, their segment shifting experiments generally do not directly investigate

the differences between EM and IM models. In fact, many IM models are capable of explaining the

segment shifting effect, as will be discussed in the concluding chapter.

3.2.2 Evidence from English suffixes

Feldman et al. (1995) have found that English suffixes are more separable from the source word

than matched pseudosuffixes. However, their design was confounded by the fact that all of their

complex words contained an embedded orthographic word while only some of their

morphologically simple words contained an embedded orthographic word (Feldman et al. 1995).

The examples in (45) are from Feldman et al. (1995). Note that all of the complex words in (45)

have embedded orthographic words while only (45a-b) of the simple words do; the words in (45c-d) do not.

		Complex	Embedded word	Simple	Embedded word
(45)	a.	winning	win	inning	inn
	b.	mobster	mob	lobster	lob
	c.	harden	hard	garden	*gard
	d.	silky	silk	city	*cit

Feldman et al. acknowledged this problem and addressed it by conducting a post hoc analysis which compared the difference in shifting times between simple and complex words which both had embedded orthographic words to the difference when only the complex words had embedded words. They found no difference between these two contrasts. They concluded that the facilitative effect found in the complex condition could not be attributed to the presence of an embedded word. However, post hoc analyses are inherently problematic and unreliable. They generally should be followed by explicit manipulations that remove the prior confound. Since it will not be possible, in subsequent experiments in this dissertation, to control for embedded words, it is crucial that I demonstrate that this confound is not responsible for the reported differences in shifting times. Thus, I will conduct a suffix shifting study analogous to prior studies conducted by Feldman and colleagues in which I contrast complex suffixed words to simple pseudosuffixed words. Since complex suffixed words almost always have embedded words, I will include only morphologically simple words that also include an embedded orthographic word. This control will make an explicit investigation of the influence an embedded word has on shifting times. This experiment will also serve to demonstrate that I can replicate Feldman's basic finding using her task. As mentioned above, ensuring that the embedded word plays no significant role in driving the results for the segment shifting task is especially important for the forthcoming investigation of free and bound root prefixed words since most bound root words do not have embedded orthographic words.

3.3 Experiment 1

Experiment 1 was conducted partly to establish a procedure for conducting a segment shifting experiment and partly to demonstrate that the presence of an embedded word is not responsible for the differences in suffix shifting times between complex and simple words. All the words in this study contain an embedded word.

Given the prior set of results obtained with this task, I expect morphologically complex words to facilitate suffix shifting relative to morphologically simple words. However, if prior facilitation effects were the result of embedded words only being consistently present in the complex condition, then I expect no difference in shifting times between the complex and simple conditions, as they both consistently include embedded orthographic words in my experiment.

3.3.1 Method

Participants. 28 native English speaking undergraduates from the University at Buffalo received partial course credit for their participation.

Materials. 34 pairs of source words which shared the same orthographic and phonological ending but which differed in their morphological complexity were paired with a single target word, such as in example (46).

(46)	Morphologically Complex	Morphologically Simple
Source word	SING<u>ER</u>	HUNG<u>ER</u>
Target word	BUY	BUY
Response word	BUY<u>ER</u>	BUY<u>ER</u>

All words in both conditions contained embedded orthographic words. Simple and complex source words were matched for stress pattern, number of syllables, log frequency of occurrence, string length, and the difference in log frequencies of the penultimate and ultimate syllables. The values for these statistics are presented in Table 3.

	Frequency	String Length	# of Syllables	Δ Syll Freq
Complex	.79 (.53)	6.7 (3)	2 (.65)	70 (99)
Simple	1.0 (.65)	6.1 (1)	2 (.17)	119 (163)

Table 3. Means (and standard deviations) for log frequency, string length, number of syllables and the difference in syllable frequencies per condition.

Items were distributed across two presentation lists such that only one member of a source word pair occurred on any list. No source or target word was repeated on either list and each participant saw only one list. A complete set of source words that were included in the analysis is given in Appendix A.

Procedure. This study used a version of Feldman & Fowler's (1987) segment shifting task. Participants were presented with morphologically complex and morphologically simple source words (SINGER vs. HUNGER). All items were presented in uppercase letters centered on the computer screen. The background color for the screen was black and the letters were presented in a light grey. Source words were presented for 750 ms at which point the (pseudo)suffix was illuminated (i.e., the display color for the suffix was changed from light grey to white) concomitant with the presentation of the target word five lines below the source word. Timing began with the presentation of the target word and ended with the participant's verbal response. In the procedure described in Feldman et al. (1995) the highlighted source word and target word are only displayed for 1500 ms at which point the screen returns blank. However, it was not clear from the description of the procedure used in prior studies whether the timer was also limited to 1500 ms or whether participants could still respond after the source and target words had been removed (Feldman 1991, 1994) It was also not clear why this time limit was imposed. Thus, in this experiment the highlighted source words and target words remained on the computer screen until a response was detected.

An experimenter remained in the room during the experiment to determine whether participants produced the correct response words and to monitor for microphone errors. Microphone errors were caused by stuttering, producing a click noise such that the microphone was prematurely tripped, and by speaking too quietly such that the microphone did not trip.

3.3.2 Results

Means and standard deviations of suffix shifting times from source words to the response word were calculated for each participant and for each item. Outliers, response times that were more extreme than 2.5 standard deviations from the participant's mean response time, were replaced by a boundary value, the mean plus 2.5 standard deviations[14]. Trials which had been marked as microphone errors were excluded from the analysis. Microphone errors accounted for 9% of all responses. Outliers accounted for 2.5% of all responses. Errors were equally distributed across conditions; thus, no error analysis was conducted. There were 48 errors in the complex condition and 42 errors in the simple condition.

Means in the complex and simple conditions were 1009 ms and 1050 ms respectively. Shifting latencies were 41 ms faster to response words formed with a morphemic string than to response words formed from a non-morphemic string. This pattern of results is represented in Figure 3.

---

[14]This outlier procedure is different from the one used in all subsequent experiments. Outliers in subsequent experiments are removed from the participant mean, not replaced by a boundary value. The present outlier procedure was applied because outliers were unevenly distributed across conditions. The simple condition had 17 outliers while the complex condition had only 8.
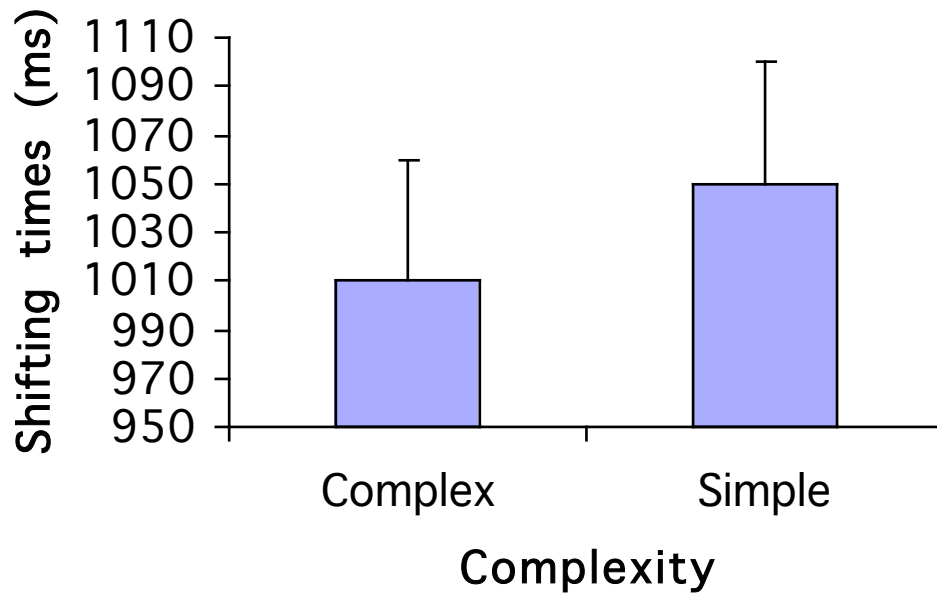
Figure 3. Suffix shifting times to target words following morphologically complex and simple source words

Participant and item means were submitted to a 2 (list) x 2 (complexity) ANOVA. The difference between experimental conditions was significant by participants, F_1 (1, 26) = 5.451, *MSE* = 4265, *p* = .02, and trended by items, F_2 (1, 32) = 2.96, *MSE* = 17072, *p* = .09.

3.3.3 Discussion

Participants found it easier to segment and shift a string when it corresponded to an actual morpheme than when it did not. Response words were produced significantly faster when affixes were shifted from morphologically complex words (SINGER) than when pseudoaffixes were shifted from morphologically simple words (HUNGER). This finding is consistent with prior findings using this task (e.g., Feldman et al. 1995); however, unlike prior studies, this study did not have the confound of embedded words consistently being present in only one condition. Thus, the obtained differences in shifting time cannot be attributed to the influence of embedded words. Furthermore, this result suggests that the segment shifting task is unaffected by the changes in the procedure that I

introduced. Allowing participants unlimited time to respond did not affect this task's ability to discriminate between morphemic and non-morphemic strings. While shifting times in this experiment are longer than the shifting times reported by Feldman et al. (1995, their mean shifting times were around 600 ms), the direction of the difference is equivalent and the magnitude of the effect is proportionate (41 ms in the current experiment and 15 ms in Experiment 1 of Feldman et al. 1995).

3.4 Segment Shifting and Prefixes

Establishing that the segment shifting task is sensitive to suffixes does not necessarily entail that it will also be sensitive to prefixes. As discussed in section 1.5.1, prefixes are different from suffixes in a number of ways, and these differences might render their structure undetectable in the task. For one thing, prefixes camouflage the beginning of the base morpheme, potentially complicating recognition. Since the beginning of the base morpheme is obscured by the prefix, the morpheme boundary between the base and the prefix may not be as easily identifiable as it is between a base and a suffix. Second, while suffixes can be either derivational or inflectional in English, prefixes are always derivational. Feldman found evidence that the segment shifting task treated inflectional and derivational suffixes differently in Serbo-Croatian (1991, 1994). She found facilitated shifting times only with inflectional suffixes; derivational suffixes did not speed shifting times in Serbo-Croatian. Since English has a very small set of inflectional suffixes, it was impossible for her to conduct a useful post hoc comparison of the behavior of English inflectional and derivational suffixes. Thus, it is possible that the segment shifting task is insensitive or less sensitive to derivational affixes than to inflectional affixes. Experiments composed solely of derivational prefixes may elicit shifting time differences that are severely attenuated. Finally, the fact that prefixes are at the beginning of the word also introduces the possibility that participants will develop a strategy in which they anticipate the production of the prefix, speeding the naming of

response words in all conditions. Since participants will always shift some initial portion of the word, they could anticipate the necessary articulation routine for producing the initial string of the response word even before the target word is presented. If participants do, in fact, develop such a strategy, differences in shifting time due to the segmentability of the source word would be lost. The effect would be reduced, masking any differences between complex and simple words. Anticipating the articulation of the response word was not a concern for suffix shifting since suffixes follow the stem. Participants cannot plan the production of the response word when shifting a suffix until they have been presented with the target word.

Clearly the usefulness of the segment shifting task for discriminating between complex and simple suffixed words does not necessarily translate into its usefulness in discriminating between complex and simple prefixed words. Any of the above considerations might render the segment shifting task ineffective for identifying complexity in prefixed words. To determine whether or not the segment shifting task is sensitive to differences in morphological structure in prefixed words, Experiment 2 will compare the shifting times to response words following three types of source word: morphologically complex prefixed words derived from free stems, morphologically complex prefixed words derived from bound roots and morphologically simple words with pseudoprefixes. This contrast should reveal initial evidence for complexity in free and bound root words, which is a central goal of this thesis.

Chapter 4

BOUND ROOTS AND FREE STEMS

This chapter presents the first prefix shifting study. This study addresses the question of whether prefixed words derived from free stems are represented as morphologically complex. It also investigates whether prefixed words derived from bound roots are represented as morphologically complex. To investigate complexity in prefixed words, I will contrast the shifting time for free stem words, such as the example in (47a) and bound root words, such as the example in (47b) to that of morphologically simple words with pseudoprefixes, such as the example in (47c). This comparison should also reveal whether the segment shifting task is sensitive to the morphological structure of prefixed words. As mentioned in section 1.5.1, the unique characteristics of prefixes may render the structure of prefixed words undetectable in the segment shifting task.

(47) a.   resell
     b.   receive
     c.   remorse

To anticipate, free stem words with prefixes, like suffixed words, will be shown to shift faster than morphologically simple control words. Evidence for complexity in bound root words, however, will be difficult to interpret. To aid the interpretation of the results for bound root words, I propose two features that may influence the way in which a bound root word is represented in the mental lexicon; namely, phonological alternation of the root and semantic transparency of the prefix. Each of these features is subsequently investigated with a post hoc analysis to determine whether they interacted with root type to influence prefix shifting time.

4.1 Experiment 2: Prefix Shifting

One way to address the question of whether words like *precede* and *receive* are represented in the mental lexicon with internal morphological structure is to compare their pattern of response

times in the segment shifting task to the pattern of response times for other types of words. For instance, prefixed words derived from bound roots could be contrasted with morphologically simple words, which are not internally structured, or other prefixed words derived from free stems, for which there is more behavioral evidence from other tasks to suggest that they are internally structured in the mental lexicon. This contrast will also demonstrate whether the segment shifting task will be useful for further investigations of prefixed words.

Experiment 2 contrasts three morphological conditions: morphologically complex words derived from free stems, morphologically complex words derived from bound roots, and morphologically simple words with pseudoprefixes. If the segment shifting task is sensitive to morphological structure in the lexical representations of prefixed words, as it has been shown to be for suffixed words, then I expect prefixed words to exhibit a pattern of results analogous to that of suffixed words. The presence of morphological structure in a source word should facilitate the production of the response word relative to when there is no morphological structure in the source word. I expect faster shifting times in the free stem condition than the morphologically simple condition. The more interesting question is how the bound root condition will behave. If speakers represent words derived from a bound root as complex, then they should have shifting times equivalent to the free stem words and faster than the simple words. If speakers do not represent words derived from a bound root as complex, then shifting times should pattern with the morphologically simple condition.

4.1.1 Method

Participants. 50 native English speaking undergraduates from the University at Buffalo received partial course credit for their participation.

Materials. 24 source word triplets were designed such that all words of a triplet shared the same orthographic and phonological word beginning but differed in their morphological complexity.

Each triplet consisted of a free stem word, a bound root word and a simple word. Source word triplets were paired with target word triples — three words that can be combined with the source word's (pseudo)prefix. An example of a source word triplet combined with a target word triplet is provided in (48).

(48)	Free stem	Bound root	Simple
Source word	DEFROST	DECEIVE	DENIAL
Target word	PRESS	COMPOSE	GRADE
Response word	DEPRESS	DECOMPOSE	DEGRADE

Three presentation lists were constructed. Each list contained all source words and all target words in unique combinations. For example, one list might have the source word- target word pairing presented in (48) while another list would have a different pairing, such as the one in (49). Crucially, when comparing the shifting time for source word triplets, the comparison was always the time to produce the same word. Each participant saw only one list and no source or target was repeated on any list. The presentation lists differed only in which source word was paired with a given target word.

(49)	Free stem	Bound root	Simple
Source word	DEFROST	DECEIVE	DENIAL
Target word	GRADE	PRESS	COMPOSE
Response word	DEGRADE	DEPRESS	DECOMPOSE

Source word triplets were roughly matched for the log frequency of occurrence, string length, number of syllables and the difference in log frequencies of the first and second syllables, but an exact match was not possible. The condition means for each of these factors are presented in Table 4. A full set of source words used in this experiment are provided in Appendix B.

	Frequency	String Length	# of Syllables	Δ Syll Freq
bound root	1.0 (.77)	6.8 (.93)	2.1 (.34)	123 (118)
free stem	.64 (.59)	8.2 (1.6)	2.8 (.78)	50 (73)
simple	1.1 (.52)	7.4 (1.3)	2.6 (.56)	75 (76)

Table 4. Means (and standard deviations) for log frequency, string length, number of syllables and the difference in syllable frequencies per condition.

The items in each condition are imperfectly matched for these three features. Free stem words are lower in frequency than the other two conditions but longer in string length and syllable number. We know that high frequency words are recognized faster than low frequency words (cf., Forster & Chambers 1973), and that shorter words are recognized faster than longer words (cf., Whaley 1978). While these mismatches are slight, they might speed the recognition of the bound root and simple conditions relative to the free stem condition, which in turn might speed the shifting times for the former two conditions relative to the latter. However, another possibility is that more frequent words have more established and unit-like representations than lower frequency words, which are more reliant on their sublexical constituents (Grossberg, Boardman & Cohen 1997). Under this hypothesis, the more frequent words would be less likely to exhibit morphological effects. Given these conflicting predictions of frequency differences, I make an attempt to keep the conditions relatively equated. The differences in these conditions are minor enough that it is unlikely they will greatly effect the results of the experiment.

Source words were also matched for other features. With four exceptions, the syllable boundary at the prefix was the same across conditions, producing either all open or all closed initial syllables across a source word triplet. The majority of the words in the two complex conditions had second syllable stress. This is a fairly common stress pattern for prefixed words in English but fairly uncommon for morphologically simple words. About half of the morphologically simple words had initial syllable stress. Much research has been conducted to

show that words with initial stress are easier to identify (in auditory presentations) than words with non-initial stress (Cutler & Carter 1987, Cutler & Norris 1988). Thus, having morphologically simple words with initial stress should only speed their recognition relative to the complex words. Additionally, eight out of the twenty-four free-stem words had a (second, optional) phonetic realization of the prefix that was different from the phonetic realization of the same string in the other conditions. For example, the word *defrost* has two possible pronunciations; one with a reduced schwa-like vowel in the initial syllable and one with a full vowel, [i]. This alternation in vowel quality is often due to the possibility of secondary stress on the initial syllable. The complications introduced by this phonological imbalance across a source word triplet will be further investigated in the discussion section.

Procedure. The procedure for this study was the same as for Experiment 1 with two exceptions. First, a microphone calibration routine was included to ensure that microphone sensitivity was optimized for each individual participant's volume level. This calibration routine consisted of setting a volume level and having the participant read a series of words that begin with a variety of segment types (stops, fricatives, glides, etc). At the end of this routine, the volume level can be accepted, in which case the participant proceeds with the experiment, or it can be modified and the word trials repeated. This calibration routine allowed for much more fine-grained control of the microphone sensitivity than was available for Experiment 1, allowing participants to speak at a more natural volume.

Second, to ensure that participants were actually reading the source word, a memory task was included after 33% of the experimental trials (24 of the 72 trials included a memory word). After participants completed an experimental trial by producing the response word, they were presented with another word and asked whether it had been the source word (i.e., whether it had been the first word of the trial). Memory words were always either the response word, the source

word or a derivational relative of the source word (e.g., the memory word COOPERATION for the source word COOPERATE). This memory task should prevent or at least reduce the participants' ability to develop a strategy of anticipating which portion of the word will be highlighted without actually reading the source word.

4.1.2 Results

Means and standard deviations for shifting times were calculated for each participant and each item. Response times that were more extreme than 2.5 standard deviations from a participant's mean response were removed from the analysis. Outliers accounted for fewer than 3% of all responses. Microphone errors accounted for fewer than 8% of all responses. Errors and outliers were equally distributed across lists and conditions; therefore, no error analysis will be conducted. There were 85 errors and 33 outliers in the simple condition, 98 errors and 28 outliers in the bound root condition, and 84 errors and 28 outliers in the free stem condition.

The mean shifting time for the free stem condition was 851 ms. The mean shifting time for the bound root condition was 864. The mean shifting time for the morphologically simple condition was 873 ms. Shifting times were faster in the free stem condition by 22 ms compared to the morphologically simple condition and by 13 ms compared to the bound root condition. Shifting times in the bound root condition were only 9 ms faster than in the morphologically simple condition. This pattern of results is represented in Figure 4.
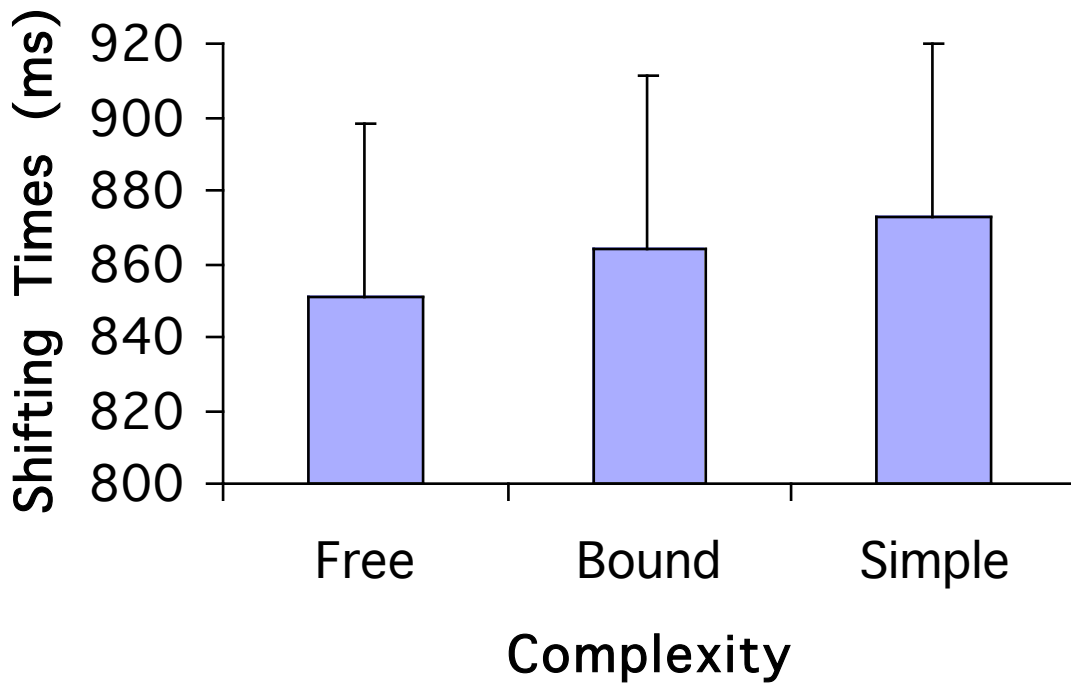
Figure 4. Shifting times to target words following bound root, free stem and morphologically simple source words

Participant and item means were submitted to a 3 (list) x 3 (complexity) ANOVA. The effect of complexity was marginally significant across the three conditions by participants, F_1 (2, 94) = 2.854, MSE= 2153, p=.06, but not by items, $F_2 < 1$. There was no effect of list, F_1 & $F_2 <$ 1.2, nor was there an interaction of complexity with list, F_1 & $F_2 < 1.6$. Three planned comparisons were performed to compare the shifting times between two conditions. The 22 ms difference between the free stem condition and the morphologically simple condition was significant only in the analysis by participants, F_1 (1, 47) = 6.28, MSE = 1946, p = .01; $F_2 < 1$. The difference between the bound root and simple conditions was not significant, F_1 & $F_2 < 1$, and neither was the difference between the bound root and free stem conditions, F_1 & $F_2 < 1.95$.

As predicted, response times to targets following complex source words with free stems were significantly faster than after morphologically simple source words, replicating the pattern of

results previously found for suffixes[15]. However, this analysis presents no evidence that the morphologically complex words with bound roots are represented any differently than the morphologically simple words. I will return to the issue of the representations for bound roots shortly. But before I do, I will first consider alternative explanations for the faster response times found in the free stem condition.

4.1.2.1 *Performance on the memory task*

As was mentioned in the procedure section, this experiment included a memory task to ensure that participants were attending to the source word. Due to the automaticity of reading and the relative familiarity of most of the source words, it is unlikely that participants were actually able to ignore the source words. However, since it is possible that some participants were not properly conducting the task, I reviewed the results of the memory task to see if any participants performed particularly poorly on it. Surprisingly, participants were generally very bad at this task, despite its presumed simplicity. Several participants (17) missed over a third of the memory questions and 11 missed over 50%. Since half of these items required a 'yes' judgment, missing 50% of the questions was potentially the equivalent of ignoring the task and responding 'yes' to all questions. To see whether performance on this task influenced participants performance on the segment shifting task, I removed those participants who missed 50% or more of the memory questions. I recalculated condition means and resubmitted them to an analysis of variance with the remaining 39 participants.

The mean of the free-stem condition was 866 ms. The mean in the bound root condition was 884 ms. The mean of the morphologically simple condition was 892 ms. Shifting times were faster in the free-stem condition by 26 ms compared to the morphologically simple condition and

---

[15] The fact that these results were not significant by items will be address in section 4.3.2

by 18 ms compared to the bound root condition. Shifting times in the bound root condition were only 8 ms faster than in the morphologically simple condition.

Participant means were submitted to a 3 (list) x 3 (complexity) ANOVA. The effect of complexity remained significant by participants, F_1 (2, 72) = 3.557, MSE = 1961, p = .03; F_2 < 1. There was no effect of list, F_1 & F_2 < 1, nor did complexity interact with list, F_1 & F_2 < 1. Three planned comparisons were performed to compare the shifting times between two conditions. The difference between the free stem condition and the morphologically simple condition remained significant in the analysis by participants, but not by items, F_1 (1, 36) = 7.7, MSE = 1704, p < .01; F_2 < 1. The difference between the bound root and simple conditions was not significant, F_1 & F_2 < 1. The difference between the bound root and free stem conditions was marginally significant in the participants analysis but not in the items analysis, F_1 (1, 36) = 3.09, MSE = 2160, p = .08; F_2 < 1.

For the most part, removing the eleven participants who performed poorly on the memory task did not change the pattern of results. Thus, for future experiments I will not use the memory task as a criterion for excluding participants from the analysis.

4.1.3 Discussion

As predicted, response times after complex words with free stems were significantly faster than after morphologically simple words. Prefixes from free stem words were more separable than a phonologically matched pseudoprefix. Faster shifting times for the free stem condition in this experiment is analogous to the faster shifting times for suffixes reported in Experiment 1 and reported repeatedly by Feldman and her colleagues.

This result suggests that segment shifting is sensitive not only to morpheme boundaries between inflectional affixes and a stem, but also to derivational morpheme boundaries. Recall that in her research on Serbo-Croatian, Feldman found faster shifting times only with inflectional

affixes, not with derivational affixes (Feldman 1991, 1994). Furthermore, all of her experiments investigating suffix shifting in English included a combination of inflectional and derivational suffixes (Feldman et al 1995, Stolz & Feldman 1995). Thus, it was initially unclear whether prefixes, which are always derivational in English, would exhibit reliably faster shifting times. This result provides reliable evidence that the segment shifting task can detect differences between complex and simple words when the morphemes being shifted are derivational. The fact that I have found a difference in shifting times for English derivational prefixes while Feldman failed to find an effect with Serbo-Croatian derivational suffixes may be due to language- specific differences or to the differences in shifting suffixes versus prefixes.

Additionally, in section 3.4, I suggested that prefix boundaries and suffix boundaries might not be treated equivalently by the task since prefix boundaries mask the onset of the word stem while suffix boundaries do not. The fact that prefixes did shift faster than pseudoprefixes suggests that the positional relationship between affix and base is not relevant for the segment shifting task. It also suggests that morpheme boundaries are as equally defined at stem beginnings as they are at stem endings. The fact that prefixes camouflage the beginning of the word root does not seem to make the morpheme boundary less salient or reduce the separability of the prefix.

Finally, the results also suggest that participants were not anticipating the production of the response words by preparing to articulate the prefix before the target word was presented. As suggested in section 3.3, if participants had developed such a strategy there should have been no difference in shifting times across any of the conditions.

4.1.3.1 Why the free stem condition was faster than the simple condition.

Faster shifting times in the free stem condition suggest that their representations are qualitatively different from the representations of simple words. It is unlikely that this difference

is attributable to the influence of embedded words being consistently present only in the free stem condition since prior investigations of this factor failed to find evidence that it affects shifting times (cf., Feldman et al. 1995, Experiment 1 of this dissertation). Rather, all the evidence to date suggests that what is at issue in determining segment shifting time is the segmentability of the shifted material as determined by its status as a morphological unit. Although the embedded word control studies were conducted on suffixed words, I generalize those results to the present study investigating the representations of prefixed words. As mentioned in the section 3.1.2, controlling for embedded words with prefixes was not possible since one condition consists exclusively of bound root words. Thus, I must rely on the generalization that, if the presence of an embedded word did not drive the difference in shifting times between complex and simple suffixed words, then it is unlikely to strongly affect the shifting times for prefixed words.

4.1.3.2 *Phonological mismatch between conditions*

Another possible explanation for the difference in shifting times found between the free stem words and the other two conditions is that some of the source word triplets differed in the actual pronunciation of the initial syllable. Some prefixes had two possible pronunciations when combined with free stems; one with a full vowel and one with a reduced vowel. In contrast, the vowels of the first syllable in the bound root and morphologically simple conditions were constant. Since the response words were also free stem words, many of them had optional secondary stress on the initial syllable as well. If participants accessed the pronunciation of the free stem source word with the full vowel and if they also produced the response word with a full vowel, then it is possible that the words in the free stem condition would have had an advantage over the other two conditions. Shifting a phonetic string that requires no modification

84

might be faster than shifting a phonetic string that needs modification. Thus, the faster shifting times in the free stem condition could have been due to this phonetic similarity.

If the shifting time differences are due mostly to segmentation in the source word, as Feldman and her colleagues argue and as I assume, phonetic consistency should have no more impact on the response times than semantic and morphosyntactic consistency have been shown to have. If I find no effect of the phonetic consistency between source and response words, then that will supply additional support to the claim that shifting time differences are truly due to separability of the affix and not in the compatibility of the shifted material with the target word.

To test this hypothesis, I removed the eight words with a pronunciation difference across the three conditions. This reduced the item triplets from twenty-four to sixteen. I then recalculated the participant and item means by condition. When all the items which have different initial syllable pronunciations across conditions are removed, the free stem condition still produced faster response words (840 ms) than the morphologically simple condition (874 ms). The bound root condition is still intermediate to the other two conditions (854 ms). The effect of complexity was significant by participants, F_1 (2, 94) = 5.01, $MSE = 2920$, $p < .01$, but not by items, $F_2 < 1.2$. There was no effect of list, F_1 & $F_2 < 1$, nor did list interact with complexity, F_1 & $F_2 < 1.9$. The difference between the free stem condition and the morphologically simple condition remained significant in the participants analysis, F_1 (1, 47) = 8.1, $MSE = 3565$, $p < .01$, and was marginally significant in the items analysis, F_2 (1, 45) = 3.6, $MSE = 4064$, $p = .06$. The difference between the bound root and free stem conditions was not significant F_1 & $F_2 < 1.9$. Interestingly, the difference between the bound root and simple conditions in this analysis was marginally significant in the participants analysis, F_1 (1, 47) = 3.72, $MSE = 2686$, $p = .06$, but did not trend in the items analysis, $F_2 < 1$.

This sub-analysis shows that when the words which were not perfectly matched for pronunciation across conditions were removed, shifting times in the free stem condition remained significantly faster than the morphologically simple condition. Thus, the faster response times in the free stem condition cannot be attributed to a better phonological match between response and source words in the free stem condition. The fact that removing these items did not change this result adds further support to the view that the locus of the segment shifting effect is in the segmentation of the source word and not in the concatenation of parts required for the production of the response word.

It is unlikely that the emergence of a difference between the bound root condition and the simple condition could have been due to improving the phonetic match between source and response words, since the bound root and morphologically simple conditions were initially better matched for initial syllable pronunciations. A more plausible explanation for it is that the removal of these items reduced the variability of the conditions.

4.1.3.3 The bound root condition's behavior

Experiment 2 demonstrates that the segment shifting task is sensitive to morphological structure in words composed of prefixes and free stems. But the shifting times to response words following bound root source words were not significantly faster than following a morphologically simple source word. They were also not significantly slower than the shifting times following free stem source words. This suggests, at first blush, that the lexical representations for bound root words and morphologically simple words are equivalent, i.e., they both lack morphological structure. However, it also suggests that bound root and free stem words are representationally equivalent.

In this experiment, and in the literature in general, bound root words are treated as a homogeneous set. However, it is possible that bound root words in English are not a

homogeneous set but rather form several different sets, some of which may be represented with internal structure while others may not. There are two factors that were not considered when constructing the materials for this experiment which could play a role in identifying a word as complex or simple, which in turn could influence the shifting times of prefixes from bound root words. These two factors will now be investigated in turn.

4.1.3.4 Types of bound root words

As discussed in section 1.2, bound root words are usually characterized by the absence of semantic compositionality. In fact, the roots are usually considered semantically opaque. Bound root words also tend to exhibit idiosyncratic phonological alternations. Thus, we have two ways in which bound root and free stem words differ. These characteristics provide a starting point from which to begin an investigation of the representations of bound root words. I will investigate bound root words by following up each of these factors. First, I will contrast roots that participate in a phonological alternation to those that do not. Second, I will contrast bound root items whose meanings are partially compositional due to the semantic transparency of the prefix to words that are non-compositional due to both the root and the prefix being semantically opaque. To anticipate, I do find an effect of semantic transparency.

4.2 Subanalyses of the Bound Root Words

In his discussion of the definition of the morpheme, Aronoff (1976) proposed that strings can be identified as bound roots if they alternate in an idiosyncratic manner, i.e., *ceive ~ ception*. However, not all words thought to be derived from bound roots alternate. Some roots have a constant form and some have a nonidiosyncratic alternation that is common in words not derived from bound roots. Thus, one possible explanation for the results in Experiment 2 is that there are

87

two types of bound root words, an alternating set and a non-alternating set, and these two sets are represented differently. Perhaps only one of them is represented with morphological structure.

A second possible confound in this study is the relative semantic contribution of the prefix. Many studies have proposed that no morphological effects will be found in the absence of semantic transparency (e.g., Marslen-Wilson et al. 1994). Since most bound root items have limited semantic content attributable to their component parts, this may reduce the available evidence for morphological structure. If it is possible to identify a subset of bound root words that have a higher degree of semantic transparency due to the contribution of a semantically transparent prefix, then I might find evidence for morphological complexity for that subset of bound root words. I will investigate each of these possibilities in turn.

4.2.1 Post hoc analysis 1: Phonological alternations

Aronoff (1976) argues that in the absence of transparent semantics, the presence of a formal alternation can cue the speaker into the presence of a root morpheme. Note for example that all the words in (50a) have the same alternation between [t] and [ʃ] when the suffix {-ion} is added. In contrast, the words in (50b) select a different form of the suffix (a different allomorph) and do not alternate [t] with [ʃ].

(50) a.   permit ~ permission
           remit ~ remission
           commit ~ commission
     b.   limit ~ limitation * limission
           vomit ~ vomitory* vomission

While an idiosyncratic alternation is a classic characteristic of bound roots, not all bound root words undergo idiosyncratic alternations. The words in (51), for example, have constant root forms. The roots maintain the same orthography and phonology, exhibiting only changes

consistent with a change in stress pattern, which are not idiosyncratic to the root but general properties of the suffixes.

(51) disdain ~ disdainful
insist ~ insistent ~ insistence
conform ~ conformation ~ conformity

In Experiment 2, most bound root source words were alternating, like the words in (50a) above. But some of the words were non-alternating, like the examples in (51). In the absence of a semantically transparent root, having a root that exhibits an idiosyncratic phonological alternation serves to distinguish the root from non-morphemic homophonous strings (Aronoff 1976). But words like *insist* and *resist* have neither consistent semantics nor a phonological alternation. It is possible that, in the case of the non-alternating words, nothing distinguishes a pair of words like *insist* and *resist* from morphologically simple words that just happen to have the same ending, like *balloon* and *saloon*. The non- alternating subset of bound root words may be mono-morphemic while the alternating bound root words are polymorphemic, since they have sufficient evidence to identify the root as a morpheme. The inclusion of monomorphemic non-alternating words in the bound root condition in Experiment 2 may have affected the response times for that condition. The alternating words may have shifted faster than the simple items, but the non-alternating items, which were averaged together with the alternating items, slowed down the average response times.

Interestingly, IM models which represent morphological information as emerging from a combination of form and meaning overlap predict that the more two related words have phonological overlap, the more a morphological effect will be apparent. Having a root that alternates will obscure a relationship, making morphological effects less robust (Seidenberg & McClelland 1989, Rueckl et al. 1997). In contrast, non-alternating roots will be more related across items, since the phonological relationship is maximized. A prediction of this view would

be that, if there is a difference between shifting times for alternating and non-alternating roots, the benefit should go to the non-alternating words.

To test these hypotheses, I performed a post hoc analysis on the data from Experiment 2. I divided the bound root condition into two sets, alternating and non-alternating, and recalculated the means. The items divided unevenly (18 alternating words and only 6 non- alternating), so these results should be regarded with considerable caution. Recall that the mean for the morphologically simple condition was 873 ms and the free stem condition shifting mean was 851 ms. In the main analysis, the mean for the bound root condition was 864 ms.[16] When the bound root condition was divided into two sets, the means split apart. While the alternating bound roots maintained a mean similar to the one reported above (874 ms) the non-alternating bound root times were much faster, with a mean of 832 ms. This is even faster than the free stem condition. This pattern of shifting times is depicted in Figure 5.
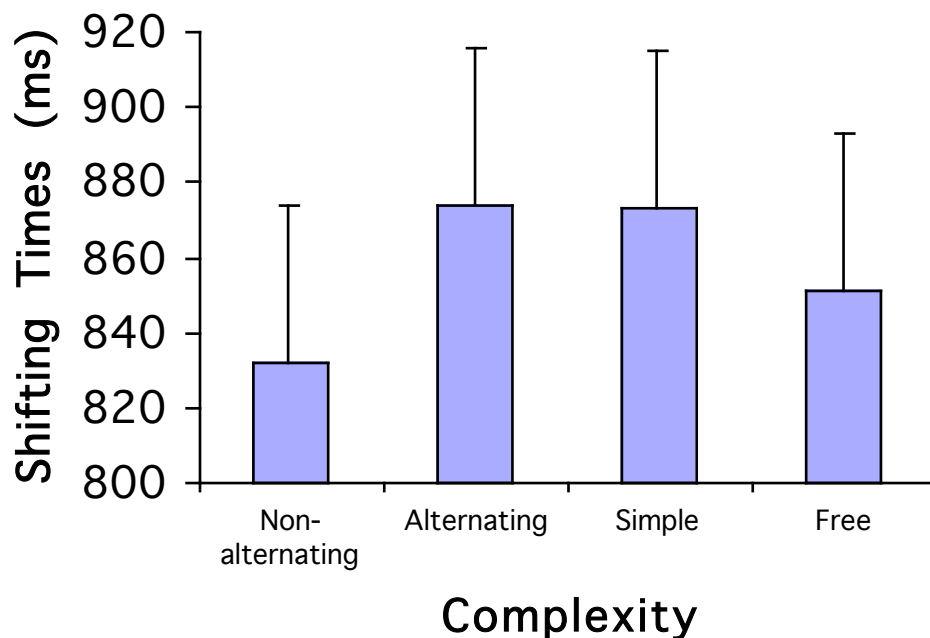
Figure 5. Shifting times to response words following simple and free stem words and bound root words that were alternating and non-alternating.

---

[16] This analysis includes the data from all 50 participants.

To determine whether these differences were statistically significant, I submitted this data to a 3 (list) by 4 (complexity) ANOVA. Because yoked item sets are divided into separate conditions, analyses will only be presented by participants, not by items. The most interesting comparison for this analysis is between the two types of bound root words. The non-alternating bound root words shifted significantly faster than the alternating words, F_1 (1, 47) = 9.35, MSE = 4636, p .01, and the morphologically simple condition, F_1 (1, 47) = 8.13, MSE = 5715, p < .01, and trended compared to the free stem words, F_1 (1, 47) = 2.9, MSE = 3336, p = .09. In contrast, the alternating items were not different from the morphologically simple condition, F_1 < 1, but they were significantly slower than the free stem condition, F_1 (1, 47) = 4.4, MSE = 2753, p < .05.

There was no benefit to having a root that alternates. In fact, non-alternating bound root words shifted faster than alternating words, which were themselves no different from simple words. This is an unexpected result given Aronoff's definition of the morpheme which relies on the presence of a phonological alternation to help identify semantically opaque bound roots. In contrast, it is not surprising from the perspective of "morpheme- less" IM models. However, it is difficult to draw reliable conclusions from a post hoc comparison in which the number of items contributing to a participant mean response time varies so greatly across conditions. Furthermore, it is important to remember that the original design of this experiment included a direct comparison of the naming times for the same response word in all three conditions. This post hoc comparison no longer maintains that control. The comparisons across the sets of words are for the production of different words with different word onsets. To make a more convincing argument regarding the benefit associated with the presence or absence of a phonological change, a study explicitly contrasting alternating and non-alternating bound root words must be conducted. That study is presented in my third experiment in Chapter 5.

To summarize, in this post hoc analysis we see a possible explanation for the varied behavior of the bound root words in the main analysis; there seem to be two very different types of words in the set. While alternating words produced shifting times that were no different from the simple words, non-alternating words shifted significantly faster than simple words. Thus, this post hoc analysis suggests that, were this factor explicitly manipulated, it would reveal an important difference in the way bound root words are represented in the lexicon.

4.3 Semantics of the Prefix

An additional confounding factor in this study is the degree of semantic transparency across prefixes. While the roots in the bound condition usually do not contribute any clear or consistent semantics, the prefixes might. The presence of a semantically transparent prefix may suffice for representing internal complexity. Comparing only those bound root items that have a semantically transparent prefix to the morphologically simple words may reveal that the bound root items shifted faster than the morphologically simple control condition. To test this hypothesis, I performed a semantic rating study on both the bound root and free stem words, to see whether speakers associate these words with the traditional and well-known meanings of the prefixes.

4.3.1 Semantic rating study 1

Most semantic transparency rating studies found in the literature are designed to gauge the transparency of a stem. To do this, two words which share a stem, such as *casualty* and *casual,* are compared to each other. In this case, however, I wanted to evaluate the semantic transparency of a prefix and I did not want it confounded with speakers' conscious awareness of the presence of a prefix or muddied by free-association or definition-elicitation tasks which provide an endless variety of responses. Therefore, I devised a semantic composition task which asks participants to rate the contribution of one word to the meaning of a second word.

I chose meanings associated with the prefix when it is transparent and paired it with words containing that prefix. This task avoids some of the limitations of prior tasks because it focuses solely on the prefix, not on the semantic transparency of the stem, and because it restricts the meanings under consideration in a way that cannot be controlled in free-association or definition-elicitation. It also maximizes usable data since no judgments are eliminated.

4.3.1.1 Method

Materials. Source words from Experiment 2 were paired with the meanings typically associated with their prefixes (Denning & Leben 1995). In cases where more than a single possible meaning is associated with the prefix, the dictionary was consulted to determine which meaning was the best fit to each word. For example, the prefix {in-} has several meanings, the most common two of which are very different. In example (52) the prefix denotes the preposition-like meaning of "in" while the prefix in the words in (53) negate the stem. It is important to pair a word such as *inscribe* with the more appropriate of the two definitions. For *inscribe*, the "in" meaning component would be a better match than the "not" meaning component.

(52) inside, indoors, inject, induce
(53) intolerable, indecent, inedible

In addition to the 24 sets of experimental items, 90 distractor items were included. Items were counterbalanced across three presentation lists and distributed in a Latin square. Example (54) shows an experimental item and a distractor item each paired with their respective meaning component.

(54) example: PRECEDE -- BEFORE
       distractor example: HISTORY -- PAST

Procedure. Item pairs were centered on the computer monitor with a 7-point Likert scale at the bottom of the screen. Participants were asked to rate the degree to which the second word contributed to the meaning of the first word. Lower ratings signify that the participants did not

identify the second word as contributing to the meaning of the first. No special attention was drawn to the prefix in the instructions or in the presentation. Instructions for this task are provided in Appendix A.

4.3.1.2 Results

Item ratings were gradient. Free stem words received the highest mean rating, 4.84 (*Standard Error* = 11), followed closely by the bound root words, 4.48 (*SE* = 11). Simple words were much lower, 3.03 (*SE* = 11). The ratings obtained from this study provide a measure of which words from Experiment 2 have a clear meaning component provided by the prefix.

4.3.2 Post Hoc analysis 2: Semantic transparency

Following several proposals from morphological theory (e.g., Nida 1946), if a word can be partially decomposed into morphemes, then any remaining string must also be a morpheme. To illustrate, consider the *cranberry* example. In *cranberry* the morpheme *berry* is easily identifiable since its form and meaning are consistent across all occurrences, like *strawberry*, *blueberry*, etc. The remaining string, *cran-* has no independent meaning, since it occurs in no other words. However, the fact that *berry* is easily identified as a morphemic unit entails that *cran-* must also be a morpheme. Words must be exhaustively parsed into morphemic units; stray phonological material that does not form part of a morpheme is generally disallowed. Thus, *cranberry* would be represented as complex in the mental lexicon. Analogously, for a word like *receive* or *precede,* if the initial string is identifiable as a prefix due to its transparent semantics, i.e., the same semantics as in fully compositional free stem examples, then the remaining string would also be considered a morpheme and the word would be represented as complex in the lexicon.

To test whether prefix shifting time was influenced by the transparency of the prefix, the materials from Experiment 2 were divided into three levels of prefix transparency based on the

score they received in the rating study. I set arbitrary rating boundaries for each condition. Any word with a mean rating below 3.99 was classified as having a semantically opaque prefix. Words with mean ratings above 5.0 were classified as having semantically transparent prefix. In an attempt to keep the categories more distinct, I also included a middle ground classification for words rated between 4.0 and 4.99. These words were excluded from post hoc analysis 2. Examples of words from each condition are provided in (55).

(55) opaque: *deformed* and *deceive*
      middle: *deduce*
      transparent: *retrace* and *predict*

Based on the transparency classification, the original items from Experiment 2 were redistributed into new conditions: transparent prefix-free stem, opaque prefix-free stem, transparent prefix-bound root, opaque prefix-bound root and two simple conditions which were divided arbitrarily into two conditions, transparent and bound.

As was the case in the phonological alternation post hoc analysis, it was not possible to compare naming times for the same word or even the same prefix because a single condition was now divided into two. Therefore, the results from this study are meant only to be suggestive of a possible explanation for the absence of a shifting time difference between the bound root condition and both the free stem and morphologically simple conditions from Experiment 2; they should be considered skeptically.

4.3.3 Results

The mean shifting times for the free stem words with semantically transparent prefixes was 849 ms while the free stem words with semantically opaque prefixes were 21 ms slower. The mean shifting times for bound root words with semantically transparent prefixes was 882 ms while bound root words with semantically opaque prefixes were 34 ms faster. The mean of the two morphologically simple condition was 872 ms, 23 ms slower than the free-transparent words,

2 ms slower than the free-opaque words, 10 ms faster than the bound-transparent words and 24 ms slower than the bound-opaque words. This pattern of shifting times is depicted in Figure 6.
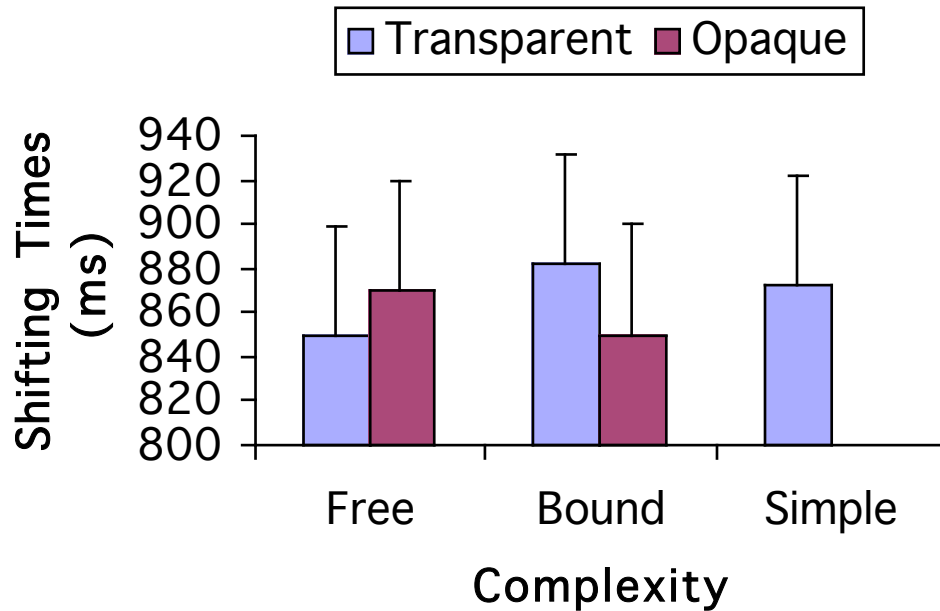
Figure 6. Shifting times to response words following free stem, bound root and simple words with both semantically opaque and transparent prefixes.

Participant means were submitted to a 3 (list) x 3 (complexity) x 2 (transparency) ANOVA.[17] The effects of complexity, list and transparency were not significant, F_1s < 1.7. The interaction of complexity with transparency was significant, F_1 (2, 94) = 3.8, MSE = 5089, p = .02. While the difference between the transparent free stem words and the simple words remained significant, F_1 (1, 47) = 6.2, MSE = 3776, p = .01, the opaque free stem words were not significantly faster than the simple condition, F_1 < 1, or significantly slower than the transparent free stem condition, F_1 < 1.4. The two bound root conditions also behave distinctly. The transparent bound root words shifted significantly slower than the opaque bound root words, F_1 (1, 47) = 7.7, MSE = 3826, p < .01. However, neither the opaque bound root condition nor the transparent bound root condition was significantly different from the morphologically simple condition, F_1s < 1.9. The simple

---

[17]As in the prior post hoc analysis, items analyses will not be presented for this data since yoked item sets are no longer maintained.

interaction of complexity and transparency for free stem and bound root words was also significant, F_1 (1, 47) = 4.9, MSE = 7659, p = .03.

The results from this post hoc analysis are provocative. While words composed of a transparent prefix and free stem mirrored the findings from the main analysis, shifting significantly faster than the simple control condition, the free stem words with opaque prefixes were not significantly differently from the simple condition. Furthermore, the bound root words with transparent prefixes were <u>slower</u> than the bound root words with opaque prefixes, not faster.

These results demonstrate that prefix transparency is an interacting factor with complexity. It appears that the transparent free stem words were carrying the main effect found in the original analysis. This may explain why the difference between free stem and simple words in the main analysis was significant by participants but not by items; half of the words had opaque prefixes and did not shift as fast as the transparent prefix words. In contrast, transparent bound root words also behave differently from the opaque bound root words, but in a different way. Instead of shifting faster than their opaque prefix counterparts, the bound root words with transparent prefixes shifted more slowly. As with the pattern found in the phonological alternation post hoc analysis, these data provide evidence that the bound root condition is not a homogeneous set of words. Rather, it seems that there is initial, tenuous evidence that both of the factors investigated with post hoc analyses interact with complexity to predict shifting times.

Since any conclusions from post hoc analyses lacking needed controls are tenuous at best, the representations and behavior of the bound root words need to be further investigated with the proper controls in place. An experiment that contrasts the shifting times of alternating and non-alternating bound root words will be presented in Chapter 5. The investigation of the role prefix transparency plays in predicting shifting times and what it can tell us about the representation of bound root words will be continued in Chapter 6.

Chapter 5

THE ROLE OF FORM

In this chapter, I investigate the hypothesis that evidence for morphological complexity can come from an idiosyncratic phonological alternation. The most common and uncontroversial sources of evidence for a morpheme is overlapping form and meaning. When two words share both of these properties, such as the words in (56), they are easily identified as sharing a morpheme. However, many words lack sufficient semantic similarity to support a morphological analysis. For example, the words in (57) share the same word-final string but do not share a clear meaning component. In the absence of shared semantics, it has been proposed that an association of a phonological unit with any formal property, such as a phonological alternation, can serve as evidence that the phonological unit is a morpheme. For example, the words of (57) both alternate in the same unusual way, as shown in (58). Furthermore, other words with similar word endings to not exhibit the same alternation pattern, as shown in (59). Thus, there must be something different in the lexical representations of words like (57) compared to the words like those in (59) to identify which words alternate and which do not. The lexical representations of the words in (57) and (59) must be distinct to account for the fact that latter do not alternate in the same way as the former, as illustrated by the ungrammatical forms in (60). The proposed difference in lexical representation between the words of (57) and the words of (59) is in their morphological structure. The former are morphologically complex, sharing a bound root which is lexically specified to take a particular form of the suffix {-ion} which licences the alternation. The latter are morphologically simple and select a different allomorph of the suffix, {-ation} or a different suffixes altogether, {-ory}.

(56) happy ~ unhappy
(57) permit ~ transmit
(58) permission ~ transmission
(59) limit ~ limitation and vomit ~ vomitory

(60) limit ~ *limission and vomit ~ *vomission

This is the general argument in favor of redefining the morpheme as a phonological string associated either with a meaning component or with some formal property. A formal property can be a phonological alteration or a grammatical function. However, as was pointed out in section 4.2, not all words which are typically identified as bound root words exhibit an alternation. For example, the entire set of {-sist} words does not alternate, as shown in (61). Nor does the set of {-sult} words, as shown in (62).

(61) a.   insist ~ insistent ~ insisting ~ insistence
     b.   resist ~ resistance ~ resistant ~ resistive ~ resistor
     c.   subsist ~ subsistence ~ subsisting
     d.   assist ~ assistance ~ assistant
     e.   consist ~ consistency ~ consistent ~ consistently
(62) a.   insult ~ insulting ~ insults ~ insulted
     b.   consult~consultant~consultation~consultative
     c.   result~ resultant ~ resulting

If the proposal that idiosyncratic phonological alternations are diagnostics for distinguishing a bound root word from a morphologically simple word, then it is possible that only words that alternate are represented as complex. The words of (61) and (62) might be represented as morphologically simple, since their roots are not associated with a consistent meaning component or a formal property. If this is the case, then perhaps evidence for the difference in their representations can be revealed by a differential pattern of segment shifting times for the alternating bound root words relative to the non-

alternating words.

5.1 Experiment 3: Phonological Alternations

As a follow up to Experiment 2's post hoc analysis 1 investigating the influence of a phonological alternation, I conducted a segment shifting study comparing words with alternating bound root to words with non-alternating bound roots. This will test more directly the argument

that a phonological alternation is a powerful cue for complexity which can stand in the absence of semantics.

Experiment 3 will provide more evidence as to whether or not a phonological alternation is a factor for identifying internal complexity, as suggested by Aronoff (1976). Three predictions are made. First, if speakers use a phonological alternation as an important cue for identifying morphemes across words, then the alternating condition should produce faster responses than either the non-alternating or the morphologically simple control condition. Second, if speakers do not use this information to help identify when a word contains a bound root, then the two bound root conditions should not be different from each other. Furthermore, following the pattern of results from Experiment 2, they may also be indistinguishable from the simple condition. To help interpret this potential null effect, this experiment also includes a free stem control condition which is predicted to be faster than the simple condition. Third, if morphological effects emerge from form and semantic overlaps, as IM models propose, then a consistent phonological pattern should enhance morphological effects while an inconsistent phonological pattern across a morphological family should reduce morphological effects. Thus, an IM model predicts that, if a difference is found between the two bound root conditions, the non- alternating condition should be faster than the alternating condition.

5.1.1 Method

Participants. 42 native English speaking undergraduates from the University at Buffalo received partial course credit for their participation.

Materials. Thirteen source word quadruplets from four conditions, alternating bound root words, non-alternating bound root words, free stem words and morphologically simple words, were paired with thirteen target word quaduplets. The crucial comparison in this study is between the bound alternating condition (e.g., *recede ~ recession, deceive ~ deception*) and the bound non-

alternating condition (e.g., *disdain ~ disdainful, retrieve ~ retrieval*). In order to facilitate the interpretation of a potential null effect, I included both the free stem condition and the morphologically simple condition as two baseline controls. A set of source words paired with one of the set of four target words is provided in (63). The complete set of source words included in the analysis is provided in Appendix C.

(63)	Alternating	Non-alternating	Free Control	Simple
Source word	REFLECT	RETRIEVE	REFUEL	RELAX
Target word	TRACE	TRACE	TRACE	TRACE
Response word	RETRACE	RETRACE	RETRACE	RETRACE

In order for a word to qualify as non-alternating, the root was required to never alternate in any word. For example, although the root {-part} does not alternate in any from of the word *impart* (e.g., *imparture, *impartion, *impartity), the fact that the same root {-part} alternates in derived forms of the word *depart* (e.g., [t] → [tʃ] in *departure*) suffices to qualify this root as alternating.

This criterion was maintained for all but two words in this study. These two words had a relative that did alternate (*involve ~ revolution* and *refuse ~ confusion*). The inclusion of two items that potentially are considered alternating may attenuate a difference between the alternating and non-alternating conditions. However, should a difference be found with the inclusion of these two items, it certainly cannot be attributed to their presence. Additionally, there was one item which was incorrectly classified as non-alternating. This item quadruplet was excluded from all analyses, leaving twelve source word sets.

Across the item sets, nine different prefixes were used, with four prefixes being repeated once. With the exception of two words with an alternating stress pattern (one in each condition), all the words in the two bound conditions had second syllable stress. All of the free stem items also had primary stress on the second syllable, however some had secondary stress on the prefix. In the simple condition, three items had initial stress; the remaining words had second syllable stress.

Syllable structure was matched across the two bound conditions in all but two instances. In the free and simple conditions there were an additional two instances in which the syllable structure did not match the morpheme boundary of the word. Other factors, such as number of syllables per word, average string length, log frequency of occurrence and the difference in log frequencies of the first and second syllables, were loosely controlled across conditions. This information is summarized in Table 5 below.

	Frequency	String Length	# of Syllables	Δ Syll Freq
Non-alternating	.94 (.62)	6.69 (.75)	2 (0)	102 (107)
Alternating	.94 (.56)	6.84 (1.0)	2 (0)	134 (120)
Free	.58 (.69)	7.4 (1.5)	2.4 (.6)	73 (74)
Simple	1.1 (.58)	6.7 (.78)	2.5 (.5)	89 (71)

Table 5: Means (and standard deviations) for log frequency, string length, number of syllables and the difference in syllable frequencies per condition.

Procedure. The procedure for this experiment was exactly the same as in Experiment 2.

5.1.2 Results

Means and standard deviations of shifting times were calculated for each participant and each item. Reaction times that were more extreme than 2.5 standard deviations from the mean for a participant were removed from the analysis. Microphone errors accounted for fewer than 13% of all responses. Outliers accounted for 3% of all responses. Outliers and errors were evenly distributed across all conditions. Both the non-alternating and the alternating conditions had 71 errors and 16 outliers. The free stem condition had 78 errors and 13 outliers and the simple condition had 68 errors and 20 outliers.

The mean shifting times in the non-alternating and alternating conditions were 852 ms and 881 ms, respectively. Shifting time for the morphologically simple condition was 865 ms, which is intermediate between the two bound conditions. The shifting times to the free-stem condition was 859 ms.[18] The pattern of response times is shown in Figure 7.
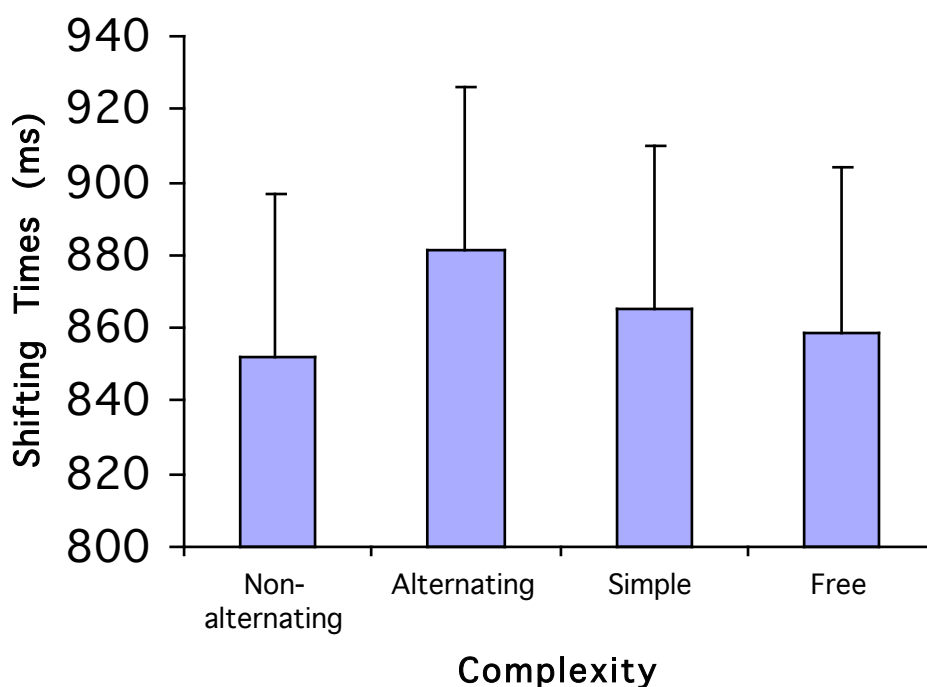
Figure 7. Shifting times to target words following morphologically simple words and morphologically complex words derived from free stems and bound roots that were either alternating or non-alternating.

---

[18] Although it is unfortunate that I have not replicated the faster response times for the free stem words here, this failure does not compromise my data for the bound root conditions. The free stem condition was not crucial to the investigation. It was only included to help interpret the potential null result between the two bound conditions and the simple condition. In the analysis of this experiment, I will not make any comparisons involving the free condition.

The free stem words may not have replicated their earlier pattern of results due to an imbalance in their surface frequency. Although the mean raw frequency of the free stem condition was matched to the other conditions, this frequency was greatly skewed by a single item, as the log frequency reveals. The mean raw frequencies in all four conditions was between 17 and 20 occurrences per million. However, without the one high frequency word, the mean frequency of the free stem condition drops to only 4.8 occurrences per million and .43 log frequency, much lower than the other conditions. If this imbalance of frequency is, in fact, responsible for the failure to replicate the faster shifting times found in Experiment 2, then it suggests that this task is not completely insensitive to factors associated with processes of word recognition.

An additional possibility is that the free stem words may have included opaque prefixes, which post hoc 2 from Experiment 2 suggested do not produce significantly faster response words than the simple condition. Unfortunately, I do not have ratings for all of the free stem words, so I cannot be sure whether this factor affected shifting times.

Participant and item means were submitted to a 4 (list) x 4 (complexity) ANOVA. The effect of complexity did not reach significance across conditions, F_1 (3, 114) = 1.51, MSE = 4026, p = .2; F_2 (3, 132) = 2.11, MSE = 5827, p = .1. There was no effect of list, F_1 & F_2 < 1, but there was a marginal interaction of list with complexity by participants, F_1 (9, 114) = 1.78, MSE = 4026, p = .08, but not by items, F_2 < 1.

The question of particular interest for this experiment is whether there was any benefit to prefix shifting for alternating vs. non-alternating bound root conditions. Planned comparisons were conducted to determine whether either of these bound root conditions was significantly different from the morphologically simple condition or different from each other. Neither the non-alternating condition, F_1 & F_2 < 1, nor the alternating condition was significantly different from the morphologically simple condition by participants, F_1 < 1.1, but the difference between the alternating and simple conditions was marginally significant by items, F_2 (1, 44) = 3.123, MSE = 7890, p = .08. Although neither of the bound root conditions was significantly different from the simple control condition, they were significantly different from each other, F_1 (1, 38) = 5.26, MSE = 3170, p = .02; F_2 (1, 44) = 4.84, MSE = 4697, p = .03.

As mentioned above, there were two items in the non-alternating condition that did not stringently meet the criteria set out in the materials section, specifically, *involve* and *refuse*. While the roots do not alternate in the source words, as shown in (64), they do alternate when combined with other prefixes (65). If words that share a bound root are all morphologically related, then a root alternation in one word might cause other word that contain the root to be treated as alternating. Thus, I compared mean shifting times to target words when these two problematic item sets were excluded from the analysis. This modification did not significantly change the general picture of shifting times (non- alternating = 853 ms, alternating = 879 ms, free stem = 854 ms, simple = 866 ms). The non-alternating and alternating bound root

conditions were still not significantly different from the simple control condition but the significant difference between the non- alternating and alternating conditions was attenuated, F_1 (1, 38) = 3.449, MSE = 4075, p = .07; F_2 (1, 36) = 3.3, MSE = 4419, p = .08.

(64) a. involve ~ involvement ~ involves ~ involving
      b. refuse ~ refusal ~ refusing ~ refused
(65) a. revolve ~ revolution ~ revolt
      b. confuse ~ confusion

5.1.3 Discussion

The aim of this experiment was to determine whether the association of a phonological unit with an idiosyncratic phonological alternation would aid in the classification of the unit as a morpheme. Aronoff (1976) argued that a morpheme need not be a meaningful unit so long as it is an identifiable unit. This proposal suggests that, in the absence of consistent semantics, roots which alternate in idiosyncratic ways may be more easily identified as morphemes than roots which do not alternate. It predicts that prefixes attached to alternating bound roots should be more separable, and in turn shift faster, than prefixes attached to non-alternating bound roots. This follows from the view that separability is, in part, determined by the relative independence of the component parts of a word. If an alternating root is more easily recognizable as a morpheme than a non-alternating root, then it should be more independent from the whole word. However, the data from Experiment 3 do not support this proposal. Not only were the alternating bound root words not different from the morphologically simple condition, they also had significantly <u>slower</u> shifting times than non-alternating bound root words. The direction of this difference suggests that, if a benefit to shifting times is to be found from the morphophonological behavior of a word, the benefit is from not alternating, i.e., maintaining a constant form across instances. This is consistent with the prediction of IM models. While it is still possible that, at some level of the linguistic system the phonological alternation plays a role

in identifying morphemes within words, there is no evidence of that in this segment shifting study.

This finding is similar to findings that have shown attenuated priming effects when the degree of phonological overlap was reduced by a single letter (Stanners et al. 1979a, Rueckl et al. 1997). However, Experiment 3 is different from prior studies in which the degree of phonological overlap was manipulated between prime and target. In this experiment, degree of phonological overlap was manipulated between the source word and other, non-presented, words in the lexicon. The fact that there was a difference between the two bound root conditions when the words that actually exhibited the phonological change were not presented, supports a view of the lexicon in which morphologically related words are connected via associative links. In other words, *retrieve* and *receive* must be lexically related to their respective derivational neighbors, *retrieval* and *reception*, in order for the differences in shared phonological form across words to have an affect on shifting times.

Since non-alternating roots are not associated to an explicit cue for morphemehood, what makes their prefixes more separable? They may be more separable than prefixes on alternating roots because of the strength of the associations between roots and their derivational relatives. When a root has a constant form, it has a clear (and stronger) association to its relatives than a root whose form changes across words. Associations to other members of a morphological family strengthens the representation of the root morpheme and in turn increases its separability from the prefix. The alternating roots have weaker relationships to their morphological family members; thus reducing the separability of the prefix.

5.1.3.1 *Investigating other contributing factors*

One possible confound in this experiment is in the relative semantic transparency of the prefixes in the two bound root conditions. The second post hoc analysis from Experiment 2

suggested prefix transparency is a potential factor for identifying morphological complexity in lexical representations. While I do not have semantic ratings for all of the words in this experiment, I do have ratings for some of them. The non-alternating condition, for which 7 of 12 have transparency ratings, has a mean prefix transparency rating of 3.79 while the alternating condition, for which all 12 items have transparency ratings, has a mean prefix transparency rating of 4.71. However, prefix transparency and shifting time were not significantly correlated in this experiment ($r < .1, p > .5$). Therefore, it is unlikely that this confound could have been responsible for the shifting time difference found between these two conditions.

5.1.3.2 Summary

The difference between the two bound root conditions suggests that a phonological alternation is not useful for identifying a string as a morpheme. However, it is interesting that neither of the bound conditions was significantly different from the simple control condition. This null difference suggests that having a constant representation across a morphological family is also not a defining characteristic for morphological complexity, although the null difference could also be due to insufficient item and participant power to reveal a small effect. This makes sense since morphologically simple words also have constant forms across their morphological family as examples (66) demonstrate. We would not want to claim that words whose forms do not change across a morphological paradigm are necessarily morphological complex. Thus, it appears that phonological constancy simply affects the strength or salience of a morphological relationship, but it is not a diagnostic for complexity.

(66) a. exam, examine, examination, examined, examiner
     b. injure, injury, injured, injurious, uninjured, reinjured
     c. delight, delightful, delightfully, delighted, delightedly

Despite the absence of a significant correlation between transparency rating and shifting times, this confound makes it impossible to determine to what extent the difference between the

two bound root conditions is due to the phonological constancy of the root versus the contribution of the semantic transparency. The only thing that seems clear from this experiment is that having a root that exhibits a phonological alternation does not help identify that root as a morpheme compared with words that do not alternate.

I predicted that if the bound root words were qualitatively different from morphologically simple words in their representations, then they should have produced statistically different shifting times compared to the simple condition. This prediction was not supported. Unfortunately, if phonological constancy is not a defining characteristic for morphological complexity, we are still left with the question of whether words composed of a prefix and bound root are represented as complex in the mental lexicon or not. It appears that the investigation of morphophonological behavior is not going to help answer this question. In the next chapter, I investigate the contribution of the semantic transparency of the prefix to the representation of bound root and free stem words.

Chapter 6

THE ROLE OF MEANING

In this chapter I investigate the contribution of a semantically transparent prefix to the representation of morphological complexity for both bound root words and free stem words. Post hoc analysis 2 from Experiment 2 suggested that prefix transparency may affect shifting times for both types of complex word, but in different ways. Free stem words with transparent prefixes produced faster shifting times than simple words while free stem words with opaque prefixes did not. However, bound root words exhibited the reverse pattern. Bound root words with transparent prefixes shifted slower than bound root words with opaque prefixes. However, since post hoc analyses often leave many factors uncontrolled for, their results must be considered with caution. Thus, the investigation of the role of semantic transparency is continued in this chapter.

6.1 Semantic Rating Study 2

To generate a pool of words from which to draw my materials for subsequent segment shifting studies, another rating study identical to the one presented in section 4.3.1 was conducted.

6.1.1 Method

Participants. 43 native English speaking undergraduates from the University at Buffalo received partial course credit for their participation.

Materials. 216 words with real or pseudoprefixes were chosen for this study; 81 free-stem, 81 bound root and 54 morphologically simple words, with 19 different prefixes and pseudoprefixes. As in Semantic Rating Study 1, presented in section 4.3.1, each word was paired with the meaning typically associated with its prefix. In cases where a prefix had multiple meanings or

multiple shades of meaning, an attempt was made to choose the meaning most likely to produce a transparent rating. Morphologically simple words with pseudoprefixes were associated with a highly typical meaning for the prefix. To allow for a wider variety of prefix meanings, the American Heritage dictionary and the Oxford English Dictionary were consulted for prefix meanings in addition to the meanings provided in Denning and Leben (1995). Examples of word-meaning pairs from the three conditions, as well as the ratings elicited for each example, are provided in (67).

(67)   <u>Opaque</u>
    Bound   REFUSE — BACK   (rating 3.57)
    Free   REACT — BACK   (rating 3.43)
    Simple   REMORSE — AGAIN   (rating 2.67)
            <u>Transparent</u>
    Bound   RECEDE — BACK   (rating 6.27)
    Free   REPLACE — AGAIN   (rating 5.00)

118 filler items were also constructed. Fillers consisted of morphologically complex and simple words paired with semi-random meanings. Many of the meaning words were prepositions, to obscure the fact that many of the prefix meanings are also prepositional. Other word-meaning pairs for fillers had part-whole relationships (DOOR- WOOD), or were semantic associates (WATER - SWIM). Generally, the filler items were designed to elicit the full range of semantic ratings, as were the experimental items.

Experimental items were distributed across three presentation lists. The number of words in each condition and with each prefix was equal across lists. Additionally, the presentation order of the words was the same across the three lists with respect to condition and prefix. In other words, the fourteenth item on all three presentation lists was a bound root word with the prefix {con-}. Likewise, the twenty-second item in all three lists was a free stem word with the prefix {re-}.

Procedure. The procedure was exactly the same as for the rating study from Experiment 2.

6.1.2 Results

Means were calculated for each item. Means per condition within a participant and across participants were also calculated. As with the prior rating study, ratings varied across the full range of the scale. The lowest item rating was 1.54 and the highest was 6.79. The mean rating for the free-stem condition was 4.85 ($SE = 11$), the mean rating for the bound root condition was 4.45 ($SE = 12$) and the mean rating for the morphologically simple condition was 3.33 ($SE = 10$). These ratings are consistent with the condition means from Semantic Rating Study 1.

Since the range of transparency ratings spanned the entire scale, I set similar arbitrary cut offs for classifying a word as transparent, opaque or intermediate as in Semantic Rating Study 1. Any item with a rating of 4.19 or below was classified as having a prefix that did not contribute significantly to the overall meaning of the word. Any word with a rating of 4.91 or above was classified as having a prefix that did contribute significantly to the meaning of the whole word. Any word with a rating between 4.20 and 4.90 was classified as having a prefix that contributes somewhat to the meaning of the word, but not consistently or strongly, and therefore these words were omitted from subsequent experiments. The intermediate, middle-ground, class was used to make the transparent and opaque classes categorically distinct.

This rating study supplied a large set of potential source words rated for prefix transparency. I used these ratings to develop materials for future segment shifting studies that explicitly investigated the differences in behavior of free stem and bound root words with transparent and opaque prefixes. Unfortunately, despite the fact that I rated over 200 words, I still encountered a shortage of words which met all of my stimulus requirements. This shortage restricted the number of conditions I could include in a single experiment. Therefore, I first investigated the role of transparency in free stem words and then investigated transparency in bound root words.

6.2 Experiment 4

Experiment 4 examines the influence of a semantically transparent vs. opaque prefix on the representation of free stem words. This experiment tests whether the results from post hoc analysis 2 in Experiment 2 can be replicated when materials are more systematically controlled. A second goal is to determine whether the basic finding from Experiment 2, namely that prefixes from free stem words shift faster than morphologically simple controls, is replicable.

This experiment contrasts the time required to produce a response word following free stem source words with transparent prefixes, free stem source words with opaque prefixes, and morphologically simple source words with phonologically equated pseudoprefixes. Competing hypotheses regarding the nature of morphological information predict three patterns of results. First, if speakers only represent internal morphological complexity when a word is fully compositional, then speeded shifting times are only expected in the transparent prefix condition relative to the morphologically simple control condition. The opaque prefix condition, which is only partially compositional, should not be different from the simple condition, under this hypothesis. I will refer to this proposal as the **full compositionality hypothesis**. A second possibility is that shifting time will be correlated with the amount of evidence that supports a complex representation. This hypothesis assumes that morphological complexity is gradient and the size of a morphological effect is determined by the amount of evidence that supports a complex analysis. Under this hypothesis, which I will refer to as the **gradient hypothesis**, the opaque prefix condition, which lacks semantic evidence for a morphemic analysis of the prefix but does have semantic evidence for a morphemic analysis of the stem, is predicted to be faster than the unstructured simple condition, but not as fast as the transparent prefix condition, which has semantic evidence for complexity from both components.[19] A third hypothesis is that

---

[19]I am assuming that all of the free stem words have a transparent relationship to their stems, although I have not

morphological information is more dichotomous or categorical in nature, being either present or not present. Following this third hypothesis, which I will call the **all-or-nothing hypothesis**, both the free stem conditions should be faster than the simple control condition since a semantically transparent free stem suffices to identify all its relatives as complex.

6.2.1 Method

Participants. 27 native English speaking undergraduates from the University at Buffalo received partial course credit for their participation.

Materials. 11 source word triplets were constructed and paired with 11 target word triplets. Each triplet consists of a free stem word with a transparent prefix, a free stem word with an opaque prefix and a morphologically simple word with a pseudoprefix. An example set of source words paired with a single target word is provided in (68). The complete set of source words used in this experiment are provided in Appendix D.

```
(68)                Transparent      Opaque          Simple
   Source word      MISCOUNT         MISCARRY        MISCHIEF
   Target word         BEHAVE           BEHAVE          BEHAVE
   Response word    MISBEHAVE        MISBEHAVE       MISBEHAVE
```

Source word sets were matched for log frequency of occurrence, string length, number of syllables and the difference in log frequencies of the first and second syllables. These measures are presented in Table 6. The match of syllable boundaries and (pseudo)morpheme boundaries were equated for all item sets. As in Experiments 2 and 3, the majority of the words had second syllable stress. However, two words in each of the complex conditions and four words in the simple condition had stressed initial syllables. Vowel quality was matched across each item set; complex words that had full vowels were matched to simple words that also had full vowels in

---

conducted a rating study to confirm it.

their pseudoprefixes. Thus, the problem of phonological imbalance that arose in Experiment 2
was eliminated in this experiment.

	Frequency[20]	String Length	# of Syllables	Δ Syll Freq
Transparent	.44 (.48)	8 (1.5)	2.6 (.67)	59 (94)
Opaque	.66 (.58)	7.9 (1.4)	2.6 (.67)	66 (102)
Simple	.55 (.38)	7.6 (1.2)	2.5 (.5)	46 (45)

Table 6. Means (and standard deviations) for log frequency, string length, number of syllables
and the difference in syllable frequencies per condition.

Items were distributed across three presentation lists. Each participant saw only one list and
each list contained all three source words of a triplet and all three targets. No source word or
target was repeated on any list. The presentation lists differed only in which source word was
paired with which target word.

Procedure. The procedure for this study is the same as for Experiment 2 with one modification
to the memory task. Memory words were never the response word itself in this experiment.
Instead they were always either one of the 3 source words, a derivational neighbor of the source
word (*preacher ~ preach)*, or a phonological neighbor (*disgust ~ discuss)*.

6.2.2 Results

Means and standard deviations for shifting times were calculated for each participant and
each item. Response times that were more extreme than 2.5 standard deviations from the overall
mean of a participant were removed from the analysis. Outliers accounted for 2% of all
responses. Microphone errors accounted for fewer than 8% of all responses. Errors and outliers
were equally distributed across lists and conditions; therefore, no error analysis will be

---

[20] Generally the words in this experiment were very low frequency. However, a shortage of words with the same
prefixes in all three condition forced the inclusion of one high frequency word in the opaque prefix condition. If this
one word were excluded from the calculation of mean frequency, then the mean token frequency for the opaque
prefix condition would also be 5 tokens per million.

conducted.   The transparent prefix condition had 22 errors and 5 outliers, the opaque prefix

condition had 25 errors and outliers and the simple condition had 23 errors and 4 outliers.

The mean shifting time for the transparent prefix condition was 797 ms. The mean shifting time

for the opaque prefix condition was 810 ms. The mean shifting time for the morphologically

simple condition was 823 ms.   Figure 8 depicts these results.
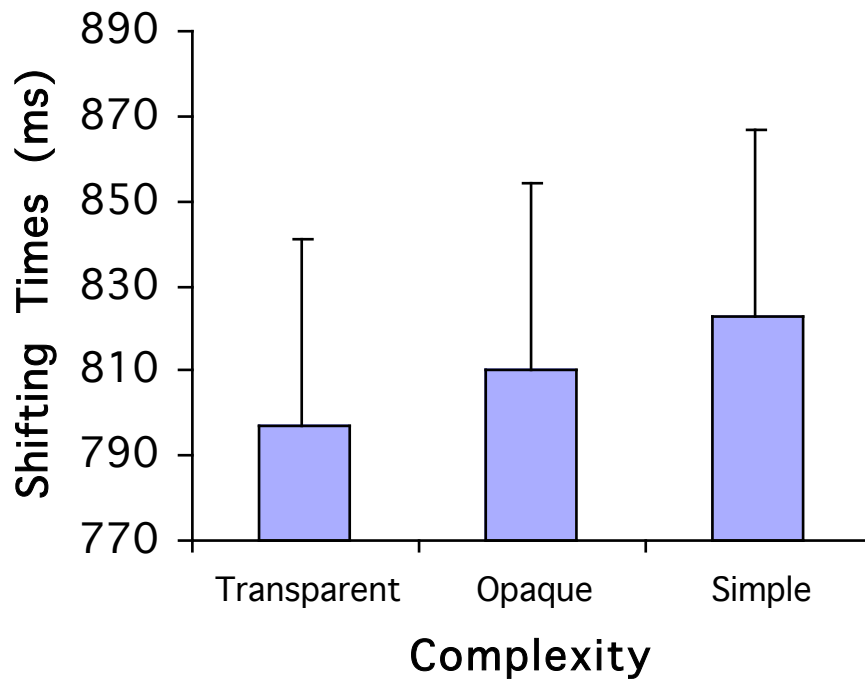
Figure 8. Shifting times to response words following free stem source words with transparent
and opaque prefixes and morphologically simple source words.

   Participant and item means were submitted to a 3 (list) x 3 (complexity) ANOVA.   Neither

the effect of complexity or list was significant across the three conditions, F_1 & $F_2 < 2.1$. Three

planned comparisons were performed to compare shifting times between two conditions. The

difference between the transparent prefix condition and the morphologically simple condition

was significant by participants, $F_1 (1, 24) = 4.93$, $MSE = 1840$, $p = .03$, but not by items, $F_2 <$

1.2.[21] Neither the differences between the opaque prefix and simple conditions nor the transparent prefix and opaque prefix conditions were significant, F_1 & $F_2 < 1$.

6.2.3 Discussion

As predicted, response words were produced faster following free stem source words with transparent prefixes than following morphologically simple source words. This successful replication of Experiment 2 supports the following conclusions initially suggested in section 4.1.3:

> Free stem words are represented with internal morphological complexity in the lexicon. They are distinct from unstructured monomorphemic words.
> The segment shifting task is sensitive to morphological structure in the lexical representations.
> The segment shifting task can detect differences between complex and simple words when the morphemes being shifted are derivational.
> The positional relationship between affix and base, i.e., does it precede or follow, does not seem to greatly affect the sensitivity of the segment shifting task to the separability of a morphemic unit.
> Participants do not anticipate the production of the response words despite the fact that they can infer what the initial syllable will be.

6.2.3.1 Free stem words with opaque prefixes

While transparent prefixes were shifted significantly faster than the pseudoprefixes of simple control words, the opaque prefixes were not.[22] The fact that the opaque prefix words did not facilitate the production of the response words to the same degree that the transparent prefix words did is somewhat surprising given the segment shifting task's presumed sensitivity to the presence of morphological structure in the source word. One could interpret the failure of the

---

[21]The items analysis for this contrast was not significant, probably due to insufficient item power. However, as seen previously, the numerical trends are still in the predicted direction. By items, the transparent condition was 21 ms faster than the simple condition and the opaque condition was 16 ms faster than the simple and only 5 ms slower than the transparent condition.

[22]Although the opaque prefix condition was not significantly different from either the transparent prefix or the simple condition, it is possible that this non-significant difference was due to insufficient item and participant power.

opaque condition to significantly facilitate the production of the response word as evidence that they are, in fact, not represented with internal morphological structure, but, that they are morphologically simple. It is possible that morphological information is included in a lexical representation only when a word is fully compositional.

But a full compositionality hypothesis is inconsistent with prior investigations of morphological information which have shown that full semantic compositionality is not necessary for morphological structure to be represented. Several studies have found morphological priming effects when semantic relatedness is absent (e.g., Emmorey 1989, Frost et al. 1997). Even more troubling is the evidence from nonceword lexical decision and naming studies discussed in section 2.2 (e.g., Caramazza et al. 1988, Laudanna, Cermele & Caramazza 1997). These studies investigated the representations of nonce words which have no meaning and found that nonce words composed of actual morphemes elicited longer lexical decision times and faster naming times than nonce words composed of non-morphemic strings. Likewise, Libben's (1993) study also included words without meaning, since he used nonce stems combined with real affixes. He also found effects of morphological structure in the absence of semantics. Furthermore, the nonceword effect was even found for bound roots, which are themselves semantically opaque (Wurm 2000). Since all of these studies used nonce word materials, their materials could not have been fully compositional. Given all of the prior evidence to suggest that full semantic compositionality is not required of a word for it to exhibit evidence of morphological complexity, it seems unlikely that some free stem words, which are clearly related to their stems, are not morphologically complex because their prefix is semantically opaque. Based on this evidence from prior studies, I reject the full compositionality hypothesis.

The all-or-nothing hypothesis predicted that there should be no difference between the opaque and transparent free stem conditions while the gradient hypothesis predicted that the opaque prefix condition should be slower than the transparent prefix condition but faster than the simple condition. Statistically, there was no difference between the two free stem conditions, supporting the all-or-nothing hypothesis. Numerically, however, the opaque prefix condition was intermediate to the other two conditions, supporting the gradient hypothesis.

The failure to find a statistical difference between the two free stem conditions may also be attributable, in part, to the sensitivity of the task. It was argued that the segment shifting task is sensitive to morphological structure in lexical representations and to the separability of affixes from the whole word. Arguably, having a semantically transparent affix could contribute to the separability of an affix, since it should make distinguishing a morphemic string from a non-morphemic string easier.

A model of morphological representations is needed that allows free stem words with opaque prefixes to be represented as morphologically complex without facilitating the production of the response words to the same degree that the transparent free stem words do. Additionally, a model of how the segment shifting task evaluates morphological complexity and semantic transparency is needed to fully understand the results. However, before a model of either can be developed, we must also understand how the semantic transparency of prefixes affects bound root words.

6.3 Experiment 5

Post hoc analysis 2 from Experiment 2 suggested that the set of bound root words was not homogeneous. Rather, it seemed that the semantic transparency of the prefix influenced its shifting time; bound root words with transparent prefixes shifted slower than the bound root words with opaque prefixes. Experiments 2 and 4 suggest that the shifting times of free stem

words was also sensitive to differences in prefix transparency. Unlike free stem words, for which there is strong evidence in the literature that they are represented as complex, bound root words have yielded comparably little evidence for the inclusion of morphological structure in their lexical representations. It is hoped that this experiment will provide some evidence for whether bound root words are ever represented as morphologically complex, and if they are, under what circumstances. Experiment 5 examines the influence of a semantically transparent or opaque prefix on the representation of words with bound roots. As was the case for Experiment 4, this is one of the comparisons from Experiment 2's post hoc analysis, except that proper controls for prefix and response word were maintained.

The claim that prefix transparency plays an important role in determining the morphological characteristics of bound root words is not unreasonable since a prefix may be the primary source of evidence for complexity in these words. Bound roots tend not to provide transparent semantics and as a result, words that share a bound root are often not highly related semantically. Furthermore, Experiment 3 suggested that the presence of an idiosyncratic phonological alternation did not provide sufficient evidence to aid in the identification of morphological complexity. Since phonological form is insufficient evidence for determining whether a string is a morpheme (as evidenced by the number of morphologically simple words with pseudoprefixes and pseudostems), semantics may be the cue that supports a complex analysis for bound root words. Therefore, it is possible that, in the absence of a semantically transparent root, the only source of evidence for morphological complexity comes from the prefix.

I predicted that bound root words with semantically transparent prefixes would elicit faster response times than the simple control condition. The bound root words with opaque prefixes, which have very little semantic evidence to support a morphemic analysis for either the prefix or the root, would be no different from the simple condition or perhaps intermediate to the other

two conditions. However, given the results of the post hoc analysis 2 from Experiment 2, I

alternatively predicted that bound root words with semantically transparent prefixes would elicit

longer response times than either the simple control condition or their opaque prefix

counterparts. This second prediction suggests that free stem words and bound root words are

qualitatively different in how morphological complexity is encoded in their lexical

representations or at least different in how the segment shifting task treats their complexity. To

investigate these possibilities, I conducted a segment shifting study that contrasted bound root

words with transparent and opaque prefixes and simple controls.

6.3.1 Method

Participants. 45 native English speaking undergraduates from the University at Buffalo

received partial course credit for their participation.

Materials. 13 source word triplets were constructed and paired with 13 target word triplets.

Each triplet consisted of a bound root word with a transparent prefix, a bound root word with an

opaque prefix and a morphologically simple word with a pseudoprefix. An example set of

source words paired with a single target word is provided in (69). A full set of source words used

in this experiment are provided in Appendix E.

(69) | | Transparent | Opaque | Simple |
|---|---|---|---|
| | Source word | REFLECT | RESPECT | REMORSE |
| | Target word | LIVE | LIVE | LIVE |
| | Response word | RELIVE | RELIVE | RELIVE |

Word sets were equated for log frequency of occurrence, string length, number of syllables

and the difference in log frequencies of the first and second syllables. The values for each of

these factors are provided in Table 7. With two exceptions, alignment of the syllable boundary

and the (pseudo)prefix boundary was the same across conditions. As in Experiment 2, the

majority of the words across conditions had second syllable stress, with only 4 exceptions. Two

words had the possibility of a fuller vowel quality in the prefix, one of which was due to a second possible pronunciation.

	Frequency	String Length	# of Syllables	Δ Syll Freq
Transparent	1.0 (.57)	6.6 (.5)	2.0 (0)	115 (100)
Opaque	.95 (.73)	7.0 (.9)	2.1 (.27)	120 (100)
Simple	.91 (.72)	7.2 (.9)	2.4 (.5)	95 (109)

Table 7: Means (and standard deviations) for log frequency, string length, number of syllables and the difference in syllable frequencies per condition.

Items were distributed across three presentation lists. Each participant saw only one list and each list contained all three source words of a triplet and all three targets. No source word or target was repeated on any list. The presentation lists differed only in which source word was paired with which target word.

Procedure. The procedure for this study is the same as for Experiment 4.

6.3.2 Results

Means and standard deviations for shifting times in each condition were calculated for each participant and each item. Response times that were more extreme than 2.5 standard deviations from the overall mean for a participant were removed from the analysis. Outliers accounted for fewer than 3% of all responses. Microphone errors accounted for fewer than 11% of all responses. Errors and outliers were equally distributed across lists and conditions, therefore, no error analysis will be conducted. The transparent prefix condition had 59 errors and 16 outliers, the opaque prefix condition had 68 errors and 18 outliers and the simple condition had 59 errors and 11 outliers.

The mean of the transparent prefix condition was 924 ms. The mean in the opaque prefix condition was 909 ms. The mean of the morphologically simple condition was 892 ms. Figure 9 presents these results graphically.
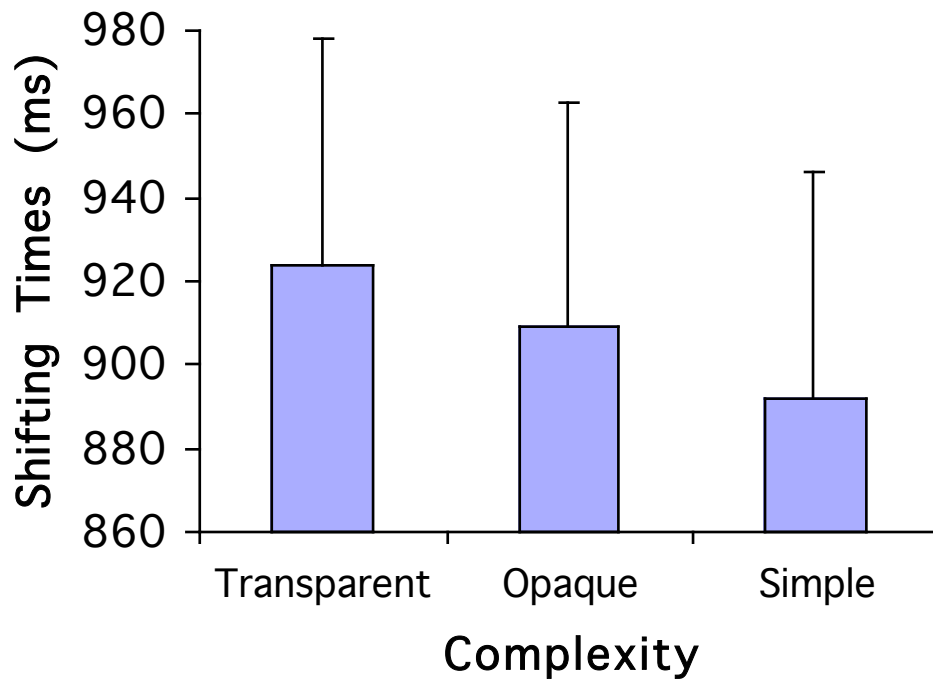
Figure 9. Shifting times to response words following bound root source words with semantically transparent and opaque prefixes and morphologically simple words.

Participant and item means were submitted to a 3 (list) x 3 (complexity) ANOVA. Complexity was the only effect to reach significance across the three conditions, but does so only in the participants analysis, F_1 (2, 84) = 3.852, MSE = 3011, p = .02, not in the items analysis, F_2 (2, 72) = .5, MSE = 11032, p > .5. Three planned comparisons were performed to compare shifting times between two conditions. The difference between the transparent prefix condition and the morphologically simple condition was significant by participants, F_1 (1, 42) = 7.516, MSE = 3078, p < .01, but not by items, F_2 (1, 36) = .96, MSE = 11540, p = .3.[23] The difference between the opaque prefix and simple conditions trended slightly by participants only,

---

[23]Generally, the differences found with the segment shifting task are quite small. There may not have been enough item power in this experiment to reveal a statistical difference. But the numerical pattern is the same in the items analysis as in the participants analysis, with a 24 ms difference between the transparent bound and the simple conditions.

F_1 (1, 42) = 2.56, *MSE* = 2659, *p* = .1; F_2 (1, 36) = .3, *MSE* = 11955, *p* > .5. The difference

between the transparent prefix and opaque prefix conditions was not significant, F_1 & F_2 <1.

6.3.3 Discussion

This study contrasted the shifting times to target words following two types of bound root

source words: those with a semantically opaque prefix and those with a semantically transparent

prefix. Source words with semantically transparent prefixes produced significantly slower

response words than morphologically simple source words with pseudoprefixes, suggesting that

the lexical representation for these two types of word are distinct. In contrast, the bound root

words with semantically opaque prefixes were numerically intermediate to the other two

conditions, but they were not significantly different from either.

This study replicates the basic finding from post hoc analysis 2 from Experiment 2 in that the

transparent bound root words shifted more slowly than the opaque bound root words. However,

the pattern of results is not parallel to that found for free stem words in Experiment 4. The

question that these results raise is why the presence of a transparent prefix slows shifting times

when it co-occurs with a bound root but facilitates shifting times when it co-occurs with a free

stem. If the segment shifting task is sensitive to the segmentability of a string, then a transparent

prefix should shift faster than a pseudoprefix, regardless of the type of root it occurs with. Bound

root words with transparent prefixes should provide more evidence for complexity than a

morphologically simple word and this evidence should facilitate shifting times.

The answer to these questions may rest in a model of morphological representations. Shifting

times for a string may depend not just on the morphological structure in a word but on whether

the word provides consistent or conflicting cues for morphological structure. Consider examples

(70)-(74), which were all rated for prefix transparency in semantic rating study 2.

(70) rebuild, encircle, prepay
(71) remodel, prochoice, subdivision
(72) react, proclaim, endear
(73) reflect, provide, inject
(74) respect, prohibit, insult

As discussed in section 1.2, the best evidence for morphological complexity is full

compositionality, such as the words in (70). Lacking full compositionality, the next best

situation is when all the morphological components of a word are semantically transparent, such

as the words in example (71). Words like (70) and (71) provide multiple sources of evidence for

morphological complexity. Each morpheme has form and meaning cues to support an analysis of

its status as a morpheme. When a word is composed of some semantically opaque morphemes,

then strong evidence for morphological complexity comes only from a sub-portion of the word.

Take, for example, the words in (72). In these words, the evidence for complexity comes only

from the stem. The morphemic status of the stem is derivable from the semantic relationship

these words hold to their respective bases. Although these words might have a clear stem

morpheme representation, the prefix itself does not provide strong evidence that it is a

morpheme, according to the ratings obtained in semantic rating study 2. Fortunately for the

words in (72), the stem provides the primary semantic content for the word. Thus, there is more

evidence for morphological complexity in (72) then in the bound root words in (73) or (74). The

words in (73) and (74) may be composed of a prefix and bound root, but in neither set of

examples is there clear semantic evidence for a root. Thus, the strongest source of evidence for

complexity is missing in these words. But at least the words of (73) have semantically

transparent prefixes. These words have some evidence for complexity, but not as much evidence

as the words of either (70) or (71). The following scale summarizes the relative amounts of

evidence for morphological complexity in the examples: 70 > 71 > 72 > 73 > 74. Note,

however, that the sheer amount of evidence for complexity is insufficient for explaining the

segment shifting results reported in Experiments 4 and 5. If segment shifting times were determined just on amount of evidence the transparent prefix bound root words should have shifted faster than simple words. Something else must be affecting the shifting times.

One possibility is that when evidence for complexity is inconsistent, the inconsistency affects the shifting time. For example, the evidence for complexity in (72) and (73) is conflicting, since one portion of the word is transparent and one portion is opaque, but it is consistent for (70) and (74). The examples in (70) have form and meaning cues to support a complex analysis from both morphemes while (74) has no evidence from meaning at all. There is no reason to posit a complex representation for words like those in (74). When a string is represented in the lexicon without an association to a morpheme, then its representation is equivalent to an unstructured simple word. In contrast, (72) has cues for complexity from the stem but not from the prefix and (73) has cues from the prefix but not from the root. Thus, examples (72) and (73) have conflicting cues; part of the word supports a complex analysis while part of the word supports a monomorphemic analysis.

The words in (70)-(73) all have some evidence for complexity. If evidence for complexity was all that mattered to the segment shifting task, then all of these words should produce faster prefix shifting times than either the words in (74) or morphologically simple words. But this is not the pattern of results that have been obtained. While words such as in (70) and (71) did shift faster than simple words like those in (74), the words in (72) were numerically faster but not significantly different and words such as those in (73) were significantly slower.

Clearly, the amount of evidence for complexity is not the only factor that determines the separability of an affix. Rather, the evidence for complexity must also be consistent; when the cues or evidence for complexity is inconsistent, segmentation is slowed. Inconsistency can come from having strong evidence for complexity from one string but weak evidence (or no evidence)

125

for complexity from another string. The segment shifting task requires participants to explicitly interrogate the representations of words for morphological structure. I argue that when the representation of morphological complexity includes conflicting, or inconsistent, cues, they interfere with the process of segment shifting. The shifting time for a string is, in part, determined by the amount of evidence and the consistency of the evidence in support of a complex analysis. The more certain a participant is that a unit is morphemic, the easier that unit is to shift. Conflicting cues introduce uncertainty about the morphemic status of the units to be shifted. As a result, participants can shift non-morphemic strings from a word with no evidence for complexity faster than morphemic units from words with inconsistent evidence for complexity. This explains why bound root words with semantically transparent prefixes produced slower shifting times than morphologically simple words.

The degree to which segmentation is slowed depends on the ratio of strong evidence for complexity and weak (or no) evidence for complexity. Evidence from a root is much more important, or weighs more, than evidence from an affix. A semantically transparent stem provides the primary source of evidence for complexity while an opaque prefix provides a weak secondary source of evidence. This inconsistency in the strength of the evidence for free stem words with opaque prefixes still produces shifting times that are faster than the morphologically simple condition, as seen in Experiment 4, because the primary source of evidence for complexity is still strong. However, the weak evidence from the opaque prefix makes the shifting time slightly slower than the free stem transparent prefix condition, in which both sources of evidence strongly supported a complex structure. Now consider the bound root conditions. A transparent prefix provides strong evidence for complexity, but it is not the primary source of evidence. The bound root, which is the primary source of evidence provides only weak evidence. The bound root words with transparent prefixes are strongly penalized for

their inconsistent evidence, resulting in even slower shifting times than the morphologically simple condition. While the bound root words with opaque prefixes were not significantly slower than the simple condition, they were numerically slower. This suggests that weak but consistent evidence is harder to evaluate than either strong cues for or against complexity. Having consistent but weak evidence interferes mildly with segmentation, producing slightly slower shifting times than were found in the morphologically simple control condition[24].

6.4 Summary

The 'consistent evidence' interpretation is driven by assumptions on how the segment shifting task calculates separability based on the representation of word complexity. The differences in the representations of complex and simple words are based on the linguistic evidence for morphological complexity. However, I have also suggested that 'inconsistent evidence' play a large role in determining the segment shifting times because of the demands that the task imposes on the participants. The segment shifting task imposes visual structure on the source word when it highlights a subportion of the word. Before segmentation can occur, participants must evaluate whether the imposed structure matches the lexical structure. When a lexical representation includes inconsistent or weak evidence, it is difficult to evaluate the fit between the imposed visual structure and the actual morphological structure in the lexical representation. It is possible that another task which does not include the same evaluative process would reveal equivalent evidence for complexity in the free stem and bound root conditions. The semantic

---

[24] Some readers might notice that the shifting times for the opaque prefix bound root condition in Experiment 5 and in the post hoc analysis 2 from Experiment 2 hold different positions relative to the simple condition. In Experiment 2 the opaque prefix bound root condition was actually faster than the simple condition while in Experiment 5 they were numerically slower. Unfortunately, it is difficult to infer anything from a comparison of results coming from different items and participants. However, the results from Experiment 2 suggest that it is possible for consistent weak cues for complexity to result in faster shifting times than no evidence for complexity, as in the morphologically simple condition. The difference in these two experiments may be do to response variability or to a systematic difference in the items selected. Furthermore, since the data from Experiment 2 was gathered from a post hoc analysis which did not maintain the appropriate controls across item sets and the differences in condition means from Experiment 5 did not reach significance, additional research is needed to determine exactly what the affect of consistent but weak evidence for complexity is on segment shifting.

factors which were so important to the segment shifting task may play a lesser role in another

task. One option is to conduct an experiment designed to elicit morphological mistakes in

speech. This will also provide converging evidence for the complexity of prefixed words from a

task other than segment shifting. This experiment will be presented in the following chapter.

Chapter 7

EVIDENCE FROM SPEECH ERRORS

In the prior chapter, I suggested that the ease of segmenting a string from the word in which it occurs is determined, in part, by the evidence in support of the string being a morphemic unit. I proposed that the segment shifting task treated semantically transparent prefixes on bound root words differently than semantically transparent prefixes on free stem words because the former provided conflicting cues for complexity while the latter provided consistent cues. The proposal was that phonological strings can be identified as morphemes if there is sufficient evidence; when sufficient evidence is provided then a strong lexical link or association is included between the lexical representation for the whole word and the lexical representation for the morphemic components. When only one morpheme in a bimorphemic word is strongly associated to its morpheme representation, then the word as a whole is said to provide conflicting cues for complexity. Separability of a string, which is calculated over the whole word, is decreased when conflicting cues are present. Since determining separability is a specific demand of the segment shifting task, the influence of conflicting cues may not be as strong in another task which does not impose morphological structure on the words presented and does not require the participant attend to morphological structure.

To supply converging evidence for my hypothesis that both bound root and free stem words are represented as complex despite the differences in shifting times when the prefixes are transparent, I present results from another methodology. I conducted a speech error elicitation task to determine whether the component parts of bound root and free stem words behave as units in a task that does not impose or explicitly manipulate morphemic material.

Before I present the study, I first present some background information on speech errors, also known as 'slips of the tongue', and prior attempts at eliciting morphological slips in laboratory environments.

A speech error is when an intended utterance is produced incorrectly. Slips come in many different types and involve many different linguistic units, such as phonetic features, phonological features, syllables, words, semantic features and morphemes. Speech errors can also involve a combination of units, such as when two words blend together resulting in an inappropriate pronunciation. Although many different linguistic units can participate in a slip, usually a single slip will involve units of a similar type. For example, syllable onsets usually exchange with other syllable onsets in the utterance and rarely with a syllable coda. Likewise, nouns usually exchange or are replaced by other nouns and rarely by verbs or adjectives and almost never by a sublexical unit.

While the majority of the morphological slips that have been reported involve inflectional morphemes, derivational morpheme slips have also been reported (see Fromkin 1973, Cutler 1980, Garrett 1980, for examples). Although morphological errors, both inflectional and derivational, do occur in natural speech, it is difficult to conclude from these that the morphological units are actually responsible for the slip, as opposed to the error being an example of a word blend or a phonological level slip. Thus, errors of the type presented in (75) might be a morphological error, due to the misapplication of a morphological process, or it might be due to a blend between two lexical competitors, such as *grouping* and *arrangement*.

		Intended	Produced	Reference
(75)	a.	grouping	groupment	Fromkin 1971
	b.	inquisitive	exquisitive	Stemberger 1985
	c.	derivation	derival	Cutler 1980

One way to differentiate between a morphological explanation and a word blend or phonological error explanation for slips that appear to involve morphemes is to investigate the

distribution of naturally and laboratory induced slips. For example, there is evidence to suggest that inflectional suffixes pattern differently in slips than other word endings, which supports a morphological interpretation of inflectional slips (MacKay 1976, Bybee and Slobin 1982, Stemberger and MacWhinney 1986). Unfortunately, the same sorts of distributional and probabilistic evidence for derivational affixes is sparse.

The existence of naturally occurring derivational errors such as in (75) have been used as evidence that roots and derivational affixes are stored and accessed separately (Fromkin 1971:45). However, alternative interpretations of derivational morpheme slips are still possible. Errors such as those in (75) could be due to word blends or phonological exchanges. To provide stronger support for a morphemic analysis of the derivational errors in (75), evidence demonstrating that morphological units participate in slips more frequently than non-morphological units is needed.   Recently, Pillion (1998) conducted a Word Order Competition (WOC; Baars and Motley 1976) study to elicit derivational morpheme errors. Prior to Pillion's use of the WOC task, this task had been used to induce spoonerisms (e.g., *darn bore* becomes *barn door,* Baars & Motley 1976), syllable exchanges (e.g., *horrible miracle* becomes *horracle mirible*, Baars, Matteson, & Cruickshank 1985) and word exchanges between phrases (e.g., He *fixed* his trousers and *dropped* his watch. Italicized words are exchanged, Baars 1977, Baars & MacKay 1978).[25] Pillion's study was designed to elicit morphological and phonological slips to determine whether morphological units were more likely to participate in speech errors than non-morpheme units.

Pillion presented pairs of French nouns and adjectives. Word pairs were either both suffixed, such as (76), or both monomorphemic words with phonologically matched pseudosuffixes, as in (77).

---

[25]In many of the prior WOC tasks the words or phrases that participants named were preceded by priming sets. Pillion did not use priming sets in her study, although I do use them in mine.

(76)	a.	troup**eau** traîn**ard**	"flock straggler"
	b.	pens**eur** haut**ain**	"thinker haughty"
	c.	croch**et** mur**al**	"hook wall"

(77)	a.	cadeau bâtard	"gift hybrid"
	b.	auteur soudain	"author unexpected"
	c.	chalet rival	"chalet rival"

Pillion presented pairs such as those in (76) and (77) to participants and instructed them that they would either have to pronounce the words in the order in which they appeared or in the reverse order. Prior studies using this task have shown that uncertainty about the order in which words are to be produced induces a higher proportion of speech errors (Baars and Motley 1976). In the case of morphological type errors, ordinal uncertainty should produce stranding errors such as in (78). The expected errors in example (78) produce one real word (the first) and one nonce word; however, Pillion also used examples that produced two nonce words, such as in (79). In each of the errors in (78) and (79), the (pseudo)roots are exchanged and the (pseudo)affixes are stranded. Pillion hypothesized that " derived word pairs should give rise, more often than control ones, to such stranding errors, if lexical retrieval actually entails the selection of intermediate morpheme-level units between words and phonemes" (1998:472).

		Presented pair	Expected slip
(78)	a.	troup**eau** traîn**ard**	traîn**eau** *troup**ard**
		"flock straggler"	"sleigh"
	b.	cadeau bâtard	bâteau *cadard
		"gift hybrid"	"boat"
(79)	a.	chat**on** frouss**ard**	*frousson *chatard
		"kitten cowardly"	
	b.	flacon standard	*standon *flacard
		"bottle standard*"*	

Pillion reported significantly more stranding errors in the complex condition than in the monomorphemic control condition. She concluded that "morpheme units like roots and derivational suffixes are handled at some stage of the real-time speech production process" (1998:487) and that derivational structure needs to be represented in the speech system.

Although laboratory-induced speech errors are often brought to bear as evidence for particular models of speech production, I would like to suggest that the task used by Pillion and, in the next experiment is no more of a production task than common word recognition tasks that require a naming response. In the WOC slips elicitation task, participants see two words on a computer screen and are required to read them aloud after they have disappeared from the screen. In most naming tasks, participants are presented with a single word on a computer screen and are required to read it aloud as quickly as possible. The only two differences between the WOC task and other naming tasks is that a) the words disappear before the participant names them and b) the participant does not know what order the words should be named. In all other respects, the WOC task and naming tasks are the same and I argue that they use the same mechanisms and representations to produce words.

Pillion's study, together with the naturally occurring derivational slips, provide compelling evidence for the inclusion of derivational morphological structure in the lexicon. It appears that the WOC procedure may be a useful mechanism for distinguishing words that include internal morphological structure in their lexical representations from words without internal morphological structure. The WOC task may be able to provide converging evidence for the results and the interpretation of the results obtained with the segment shifting task in Experiments 1-5.

The conflicting cues hypothesis proposed for the results of Experiments 4 and 5 suggested that both free stem words and bound root words with transparent prefixes are represented as complex. But when only part of the word supports a complex analysis, such as when one morpheme is semantically opaque, the representation provides conflicting cues for segment shifting. Having a portion of the word which appears to not be associated with any morpheme reduces the separability of the affix, which is what the segment shifting task is apparently

133

sensitive to. This hypothesis still maintains that bound root words can include morphological complexity so long as at least one component is sufficiently associated to its morpheme representation. A slips elicitation task manipulates word components, without explicitly imposing a visual structure on a word, as the segment shifting task does. Thus, if the prefix and/or root in the bound root words are considered morphemic by the language processor, then slips should be produced.

7.1 Experiment 6

In order to provide converging evidence for the internal structure of prefixed words from a different methodology as well as to test which hypothesis best explains the pattern of results obtained in Experiments 4 and 5, I conducted a WOC experiment contrasting the number of elicited slips in three conditions: bound root, free stem, and morphologically simple. Because it has also been shown that errors that produce actual words are more likely than errors that produce nonsense strings I only included materials that would produce real word errors (Baars, Motley & MacKay 1975). Unfortunately, the constraint against nonsense errors restricted the possible construction of morphologically simple words with matching pseudoprefixes. Therefore, unlike Pillon's monomorphemic control group which was phonetically matched to the complex group, as shown above in examples (76)-(78), many of the words in my control group consist of morphologically simple words without pseudoprefixes but whose syllables are exchangeable, such as in (80).

		Presented pair	Expected slip
(80)	a.	regal legion	legal region
	b.	rattle morbid	mortal rabid

To determine whether morphological slips of the type reported by Pillon for French suffixed words can be elicited for prefixed words, I contrasted the number of morphological slips elicited

by free stem prefixed words to the number of syllable-level slips elicited in the morphologically simple condition. If morphological information is encoded in the lexicon and used by the language processing system, more slips should be produced in the free stem condition than in the simple condition. In addition to free stem and simple word pairs, I also included a bound root condition. If bound root words with prefixes are represented with internal morphological structure, then I predict more slips in the bound root condition than in the monomorphemic condition. In contrast, if bound root words are not represented with internal structure, then I would expect no more slips in the bound root condition than are produced in the monomorphemic condition.

7.1.1 Method

Participants. 65 native English speaking undergraduates from the University at Buffalo received partial course credit for their participation.

Materials. Eleven word pairs were constructed in each of the three conditions: free stem, bound root and morphologically simple, resulting in thirty-three word pairs. Word pairs were designed such that if the beginnings of the two words were exchanged, they would still produce two real words. The word pairs that were presented to the participants were semantically and phonologically unrelated to each other. Of the 66 words that comprised the 33 pairs, most (51) were bisyllabic, and 15 were trisyllabic. Five pairs had an imbalance in number of syllables across the two words. The remaining twenty-eight pairs had the same number of syllables and the same stress pattern across a word pair.

The mean (and standard deviation) log frequencies of occurrence for word pairs (and the intended slips) for each condition was .4 (.5) tokens per million for free stem words, .88 (.6) tokens per million for bound root words, and .86 (.7) tokens per million for simple words. When there was a large frequency difference between the presented word pairs and the potential

intruders, the intruder words were always higher in frequency on the hypothesis that errors occur

when a competing word obtains higher activation than the presented word. If the presented

words had higher frequencies than the potential errors, the likelihood of an error would have

decreased.

A sample set of items is provided in (81). A full set of experimental word pairs are presented

in Appendix F. Additionally, 72 filler items were interspersed with the experimental items.

Filler items consisted mostly of random word pairs. However, they also included rhyming

words (e.g., *heel kneel*), alliterative words (e.g., *greedy green*), semantically associated words

(e.g., *yellow white*), and words standing in an unusual syntagmatic relationship (e.g., *man hairy*).

All items were distributed within a single list, resulting in a fully within-participants design.

(81) <u>Bound Prefixed</u>   <u>Possible Error</u>
    reject **in**duce   inject reduce
    <u>Free Prefixed</u>
    preheat **re**soak   reheat presoak
    <u>Simple</u>
    shallow **wi**dow   shadow willow

For the two complex conditions, the splicing point, i.e., the point at which the words would

have to be divided in order to produce real word errors, occurred at the morpheme boundary

which usually coincided with a syllable boundary. Occasionally, a consonant of the prefix

would be ambisyllabic, as in *exclaim*, resulting phonetically in a mismatch between syllable and

morpheme boundaries. One bound root pair's potential error required a different allomorph for

the prefix (*project succeed _ subject proceed*). In the morphologically simple condition, most of

the splicing points occurred at the syllable boundary. However, in one example the syllable

boundary was not respected (*fashion* and *sector*) and in two examples onset shifting was

sufficient to produce real word slips (*mallet ~ patron* and *ditches ~ sweater*). Additionally, three

of the simple word pairs' potential errors were spelled differently from the presented word pairs;

these potential errors where phonologically matched but not orthographically matched (*fashion*

sector _ factor session). Evidence from prior studies using this task suggest that conflicting orthography should not matter as long as the phonological representations are consistent (Baars & Motley 1976).

Procedure. The procedure was borrowed directly from Baars and Motley (1976) and included some elements not adopted by Pillion (1998). Experimental word pairs were preceded by three sets of priming pairs. The priming pairs were word pairs that shared the same word onsets as the experimental word pairs. The purpose of the priming pairs was to induce an order expectancy. The participant is presented with a series of word pairs. She is instructed that she will only need to respond to some of the pairs, but she will not know which pairs need a response until the words have been removed from the computer screen. Thus, she must be prepared to respond to all pairs. The priming pairs establish a pattern of expected word order. Each pair of a priming set consists of the same word beginnings, as in example (82)-(83), taken from Baars & Motely (1976).

(82)	sane foam	(83)	big ditch
	sell phone		bill deal
	seal fog		bark dog
	seem fine RESPOND		darn bore RESPOND

In example (82), the first word of each pair begins with the phoneme [s] and the second word begins with the phoneme [f]. The repeated order of [s] first followed by [f] forms an expectation for subsequent pairs. Thus, if the participant is asked to repeat the last word pair, she can do it easily, because the order is consistent with her expectation. But if the experimental word pair is inconsistent with the priming pairs, as in example (83), then the chances of an error are greatly increased (Baars & Motley 1976).

The priming word pairs in the current experiment shared the entire initial syllable with the experimental word pairs. However, because I did not want to form an expectancy for morphological complexity in the two complex conditions, not all the priming word pairs were

morphologically complex. Instead, the priming word pairs (a total of six words or three pairs) consisted of two free stem words, two bound root words and two morphologically simple words with pseudoprefixes. An example of a priming set with the bound root experimental word pair is provided in (84). The morphologically simple word pairs had priming words that were all simple.[26] An example of priming word pairs and a morphologically simple experimental word pair is given in (85). In addition to the experimental word pairs and the priming pairs that preceded them, 72 distractor items were also preceded by sets of priming pairs. To reduce the predictability of when a response would be required, distractor pairs were preceded by zero to three priming pairs.

(84) receive insane   (85) shady willing
      repair indian   shanty wilted
      remodel involve   shifty winter
      reject induce   REVERSE   shallow widow REVERSE

Prime and experimental word pairs were centered on the computer screen in capital letters. Each pair of words remained on the computer screen for one second with a 300 ms. lag separating the word pairs from one another and the response cue.[27] Participants were instructed that they would be presented with many word pairs but were only required to respond to some of them. Thus, they should be prepared to read each set aloud. Participants were also instructed that responses would be of two types; they would either have to repeat the words in their presented order or in the reverse order. The task was fast-paced and some trials were lost due to participants' inability to respond.

---

[26] The motivation for the priming word pairs in the complex condition being a mixture of different types of words was to avoid an expectancy for a particular type of word. However, since the simple control condition was not composed of words with pseudoprefixes, it was impossible to have an equal mix of prime types in this condition. This means that in the complex conditions at least two of the six prime words could be priming a morphological structure.

[27]This is a slightly longer presentation time than the 900 ms used in Baars and Motley (1976) with a 200 ms interval. This adjustment was made because the stimulus words in this study were mostly bisyllabic rather than monosyllabic.

Word pairs were removed from the screen before participants knew whether or not they needed to respond. After a word pair was removed from the screen, it was followed either by another word pair, which indicated that no response was required, or by a response cue centered on the screen to either REPEAT or REVERSE. Removing the word pairs from the monitor before the response cue was given prevented participants from simply reading the words off of the screen.

Participants sat before a computer monitor and a microphone. There was also a tape recorder in the room. Trials were recorded and then scored for speech errors by one of two raters. Errors were transcribed and then classified into different types of errors.

7.1.2 Results

7.1.2.1 Classification of errors.

The most common errors were failure to reverse the word order, failure to complete the pair, mispronunciations (various phonological slips), wrong words produced (external intrusions), inflectional errors, and morphological errors. Examples of these errors are presented in Table 8.

Intended utterance	Produced error	Error Type
reject induce	induce reject	failure to reverse the word order
dispose propel	dispose	failure to complete the pair
reject induce	reduce	
rattle morbid	battle morbid	mispronunciations (phonological slips)
remit advise	[ridm_t] advise	
widow shallow	window shallow	wrong words (external intrusions)
recite exclaim	recite refrain	
reject induce	reject reduce	Morphological errors
reject induce	reject inject	
reject induce	reduce inject	
pretrial misjudged	pretrial misjudge	inflectional errors
inaction transform	inaction transformation	other derivational errors
remit advise	readmit admit	
rewind undone	redone unrewind	

Table 8: Examples of errors organized by error type

Only the morphological errors were analyzed for this study. Of the morphological errors, there were three types: prefix spreading, where both words were produced with the same prefix, such as in (86), stem spreading, where both words shared the same stem, such as in (87), and total morpheme exchanges, where both prefix and stems were swapped, such as in (88). These three morpheme type slips were the only slips included in the analysis. In order for a speech error to be included in the analysis it had to be fully completed. Partial slips, even one that looked as though it must result in a morphological slip (*reject induce becomes reject re... or reject induce becomes reduce...),* were not included. Thus, if anything, the data represent an underestimate of the morphological errors produced.

	Correct response	Error	Error type
(85)	reject induce	reject reduce	Prefix spreading
(86)	reject induce	reject inject	Stem spreading
(87)	reject induce	reduce inject	Total exchange

7.1.2.2 Distribution of errors.

Each of the 65 participants contributed a maximum of 33 responses, creating 2145 opportunities for error, 715 opportunities per condition. Because the task was quite fast- paced, participants were occasionally unable to respond. In total, 148 trials, (7% of all the trials) were skipped by participants. Fifty-eight trials were skipped in the free stem condition, 52 trials were skipped in the bound root condition and 38 trials were skipped in the simple condition. Additionally, 96 more trials (an additional 4%) were incomplete. There were 33 incomplete trials in the free stem condition, 49 in the bound root condition, and 14 in the simple condition. Since I required full production of both words to classify an error as one of my three error types, an incomplete response did not constitute an opportunity for an error. Since more trials were skipped or incomplete in the two complex conditions, there were more completed trials in the simple condition and thus more opportunity for error in the simple condition.

The mean number of morphological errors were calculated for each participant in each condition. On average, participants made very few morphological errors. In the free stem condition participants made an average of 1.03 morphological slips. In the bound root condition participants made an average of 1.09 morphological slips. In the simple condition participants made an average of .69 morpheme-like slips. Across the 65 participants, each free stem pair elicited an average of 6.09 slips. Each bound root item had an average of 6.45 errors and each simple items had an average of 4.09 errors per item. The error patterns by participants and by items are presented in Figures 10 and 11, respectively.
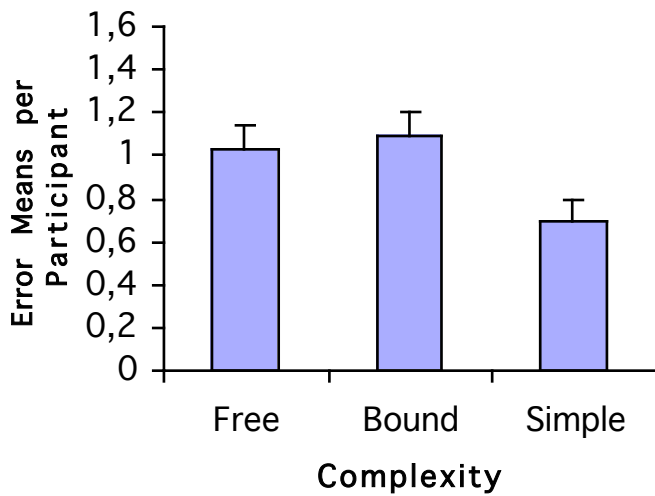
Figure 10. Mean number of slips elicited per participant by morphologically simple word pairs and morphologically complex word pairs derived either from free stems or bound roots.
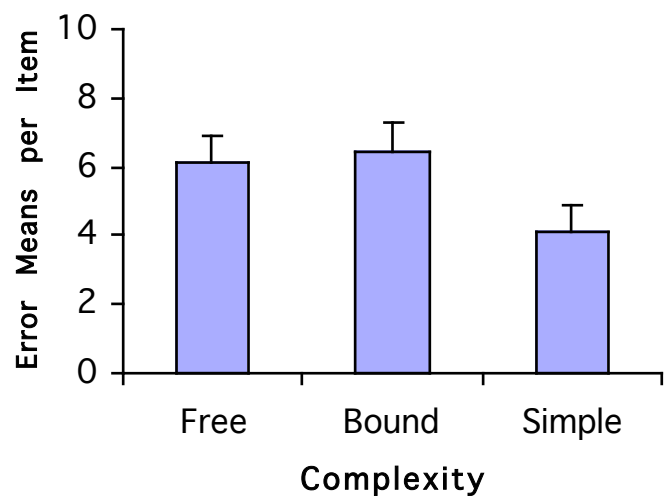
Figure 11. Mean number of slips elicited per item by morphologically simple word pairs and morphologically complex word pairs derived either from free stems or bound roots.

Participant means were submitted to a within-participants analysis of variance with 3 levels of complexity. The effect of complexity was significance across conditions, F_1 (2, 128) = 4.022, MSE = .75, p = .02. The free stem condition was not different from the bound root condition, F_1 (1, 64)=.18, MSE = .68, p > .5. However, both the bound root condition, F_1 (1, 64) =6.186, MSE = .84, p = .01, and the free stem condition, F_1 (1, 64)= 5.14, MSE = .72, p = .03, were significantly different from the morphologically simple condition. Because item sets were not paired or yoked in this experiment as they were in all the prior experiments an analysis of variance is not the appropriate statistic to conduct on the item data. Instead, I conducted three unpaired t-tests on the items data. However, none of these contrasts were significant. Unfortunately, slips data tends to be quite impoverished, with the average participant only producing two or three usable error responses. Although I had sufficient participant power to reveal a significant difference, my item power was not great enough. It should be noted that Pillion used 60 participants and 60 item sets with only two conditions.

7.1.3 Discussion

As predicted, the free stem condition produced more slips than the morphologically simple condition, even though the simple condition had greater opportunity for error given that there were fewer skipped and incomplete trials. Likewise, the bound root condition also produced as many slips as the free stem condition and more slips than the simple condition. Again, due to the disproportionate number of skipped and incomplete trials, the bound root condition had fewer opportunities to produce a slip.

As mentioned above, the morphological speech errors were scored into three types: prefix spreads, stem spreads, and total exchanges. The distribution of errors by type are in Table 9.

	Free stem	Bound root	Simple
Prefix spreads	28 retest preview → retest review	20 deflect recline → reflect recline	3 money hunky → monkey hunky
Stem spreads	13 rewind undone → unwind rewind	15 disruption contortion → distortion contortion	25 legion regal → legion region
Total exchanges	26 reappear discharge → disappear recharge	36 expose oppress → oppose express	17 rattle morbid → rabid mortal
Total errors	67	71	45

Table 9: Total number of speech errors per condition divided into error types.

While it is difficult to conduct statistical comparisons on the subtypes of errors when the numbers are so small, two striking observations emerge. First, there are no large differences between the free stem condition or the bound root condition in any type of error, suggesting that participants treated them the same way. Second, while the two complex conditions had many

143

errors involving the spreading of the first syllable (the prefix), there were almost no such errors in the morphologically simple condition. But there were more second syllable errors in the simple condition than in either of the complex conditions.[28] Thus, not only are more overall errors produced in the two complex conditions than in the simple condition, but the type of errors are different as well.

The production of a stem spread slip in the simple condition results in two words that rhyme, such as in (89). The production of a prefix spread slip in the simple condition results in two alliterative words, as in (90). The distinct distributions in number and type of speech errors produced between the two complex conditions vs. the simple condition supports the claim that the slips produced in the simple condition were different in nature from the slips elicited in the complex conditions. Thus, they support a view that the slips in the complex condition involved different units than the slips in the monomorphemic condition. Morphological units were exchanged and stranded in the former while phonological units were manipulated in the latter.

	Correct response	Error	Error type
(26)	legion regal	legion region	Stem spreading, rhyming
(27)	legion regal	legion legal	Prefix spreading, alliteration

Furthermore, the fact that there was no difference in number or type of errors in the two complex conditions suggests that the distinction between free and bound roots is irrelevant in this task. In terms of ordinal confusability, bound roots and free stems are treated the same way by the task and both are distinct from a non-morphemic phonological unit.

7.1.3.1 The role of semantics

One stark contrast between the results of Experiment 6 and the results of the segment shifting studies, specifically Experiments 4 and 5, is that the WOC study supplied no evidence for a

---

[28] The absence of first syllable slips in the simple condition is consistent with the intuition that, in English, words which rhyme are phonologically more similar than words which share word onsets. Note that the alliterative words *cat, cad, cap, cast* are less similar to each other than the rhyming words *cat, hat, mat, sat*.

distinction between the bound root and free stem words while segment shifting did. Both

conditions produced generally equivalent numbers and types of speech errors. In contrast, the

pattern of results from the segment shifting studies for the bound root and free stem words

differed. Bound root and free stem words elicited differing shifting results which were

predicted, in part, by prefix transparency. Semantics often plays a role in naturally occurring

and laboratory-induced speech errors (Fromkin 1971, Motley & Baars 1976). Thus, I was

interested in whether more errors were produced when the words of a presentation pair were

related to their intruder words relative to when they were semantically unrelated. In other words,

were more errors produced when the words of a presentation pair (e.g., *deflect recline*) were

semantically related to their intruders (e.g., *reflect decline*) relative to when the presentation

pairs (e.g., *expose oppress*) were unrelated to their intruders (e.g., *oppose express*). Semantic

relationships also provides evidence for the morphological complexity of these words.

Therefore, I was interested in whether the effect of semantic relatedness was stronger in the

bound root condition, where evidence for complexity is often lacking.

To address these questions, I conducted a semantic rating study and then redistributed errors

based on whether the presentation pair was related or unrelated to their intruders. I found no

evidence that semantic relatedness increased the likelihood of an error. There were more errors

in the unrelated free stem condition than in the related free stem condition. Furthermore, errors

were distributed evenly between the related and unrelated pairs in the bound root condition.

Thus, semantics did not seem to play a larger role in the bound root condition than in the free

stem condition.

In contrast to the segment shifting task, the WOC slips elicitation technique did not provide

the same evidence for a distinction between types of bound root words. It also did not seem to

be affected by issues of semantic relatedness. However, it does provide converging evidence

that free stem and bound root words are both composed of discrete processing units, as suggested in my interpretation of Experiments 4 and 5.

In the next chapter I conclude my investigation of the representations of prefixed word with a proposal for how morphological information is represented in the mental lexicon. My model will account for all the data presented in Experiments 1-6. Following the presentation of the model, I will discuss which of the linguistic theories of morphology are most consistent with it and propose changes which would allow the models that are not consistent with it to capture the data presented here.

Chapter 8

CONCLUSION

This dissertation has presented the results of five segment shifting studies and one speech error elicitation study. Additionally, three semantic rating studies were conducted. The results of these studies suggest a picture of lexically encoded morphological information based on lexical relationships between complex words and the representations of component morphemes.

8.1 Summary of Results

Experiment 1 was a suffix shifting study designed to replicate the basic findings reported by Feldman and her colleagues (e.g., Feldman et al. 1995), while controlling for the presence of embedded orthographic words. I replicated the typical pattern of results; morphological strings shifted faster, producing faster response words than their non- morphological controls. This result was not attributable to the influence of an embedded word. Since a primary interest of this dissertation was to investigate the representations of prefixed words derived from bound roots and free stems, this control was crucial for the interpretation of subsequent experiments.

Experiment 2 supplied the first evidence for morphological structure in lexical representations from prefix shifting. This experiment provided evidence that prefixes from free words shifted

significantly faster than pseudoprefixes from morphologically simple words. Prefixes from words derived from bound roots were numerically intermediate to the other two conditions but not significantly different from either. Two post hoc analyses were also conducted on the words in this experiment to investigate possible explanations for the failure to find speeded shifting times in the bound root condition. One post hoc analysis investigated the difference in shifting times for words composed of bound roots that alternated phonologically to those whose roots had a constant phonological form. A second post hoc analysis investigated the difference in shifting times for bound root and free stem words whose prefixes either did or did not contribute transparent semantics to the meaning of the whole word. Both of these post hoc analyses suggested that the investigated characteristics may be factors in predicting the shifting times for the prefixes. Each of the factors was followed up with an experiment in which the characteristic was explicitly manipulated.

Experiment 3 continued the investigation of prefixed words, focusing on words derived from bound roots. This experiment examined the contribution of an idiosyncratic phonological alternation to the classification of a string as a morpheme. This study confirmed the pattern of results that was suggested by post hoc analysis 1, namely prefixes from non-alternating bound root words shifted faster than prefixes from alternating bound root words. However, neither of these conditions was significantly different from the morphologically simple condition. Thus, it was suggested that neither of these features served as a defining characteristic of morphological complexity for bound root words.[29] Rather, this experiment suggested that morphological relations can vary in their strengths, based on the amount of evidence for complexity. Non-alternating roots, which have more form overlap with their morphological neighbors, had a stronger, more salient, root representation than the alternating bound root words.

---

[29]A lack of sufficient power was also proposed as a possible explanation for this null effect

Experiments 4 and 5 followed up on the findings of post hoc analysis 2 by evaluating the effect of a transparent versus opaque prefix on free stem words (Experiment 4) and bound root words (Experiment 5). These two experiments replicated the basic pattern of results from post hoc analysis 2; transparent prefixes from free stem words shifted significantly faster than morphologically simple controls. Opaque prefixes from free stem words shifted numerically faster than morphologically simple controls, but this difference was not significant. The pattern of results for bound root words was quite different. Transparent prefixes from bound root words were shifted slower than morphologically simple controls. Opaque prefixes from bound roots were numerically slower than their simple controls, but, like their free stem counterparts, this difference was not significant.

Finally, Experiment 6 provided evidence for the morphological complexity of both free stem and bound root words from a slips elicitation study. More morphological errors were elicited in both complex conditions than non-morphemic errors were elicited in the control condition. Furthermore, the types of errors elicited differed across conditions. While the simple condition elicited almost no prefix exchange errors, both complex conditions elicited slips in all three categories. Thus, Experiment 6 provides converging evidence from a different methodology for the morphological complexity of prefixed words derived from both free stems and bound roots.

As was mentioned in several of the discussion sections throughout this dissertation, the analyses by items frequently did not reach significance. I proposed a number of different explanations for this failing. For Experiment 2, the results of post hoc analysis 2 showed that the free stem words with opaque prefixes did not pattern the same as the free stem words with transparent prefixes. The fact that only half of the items were shifting significantly faster than the simple condition resulted in non-significant differences in the analysis by items. However, when prefix transparency was controlled, as in Experiments 4 and 5, the items analyses were still

not significant. I suggested that these problematic results, and the analogous one in Experiment 6, were due to a shortage of item power. All three of these experiments had many fewer items per condition than are typically used for the tasks. It is hoped that, were these experiments to be re-run with more items, significant differences by items would emerge. However, until that time, generalizations cannot be extended to the entire lexicon; interpretations must be limited only to the words included in these studies.

This cautious position does not diminish the contribution of these studies. I have shown that stem type, form constancy, and prefix transparency are three factors that affect the realization of morphological structure in the lexical representations of prefixed words. However, it is also known that factors such as prefix likelihood (Wurm 1997) and morphological family size (Schreuder & Baayen 1995) also affect the realization of morphological structure in the lexicon. Therefore, in each of my studies, these and other factors may have been influencing the results. Since the interaction of these factors is very complex, it is difficult to make broad generalizations about the nature of all words when only a subset have been investigated. Until more is known about the interaction of the factors introduced in this dissertation with the many other factors discussed in the literature, generalizations over the whole lexicon would be inappropriate. Therefore, the interpretation of my results and the following discussion should be considered an idealization and the fact that many unknowns remain should be kept in mind.

8.2 Goals and Focus

This dissertation presents an investigation of prefixed and pseudoprefixed words in English. The goal was to find evidence of morphological structure in lexical representations. The three characteristics of prefixed words focused on were: (a) the status of the root as a free or bound morpheme, (b) the phonological constancy of the root, and (c) the semantic transparency of the

prefix. The investigation was conducted with a task that is sensitive to the morphological structure of words as well as to the separability of individual morphemes from the whole word.

8.2.1 Free stem words

The results of the present studies provided evidence for morphological complexity in the representations of prefixed words derived from free stems. Additionally, the semantic transparency of a word's prefix was a factor which influenced shifting times. Only free stem words with semantically transparent prefixes elicited significantly faster prefix shifting times than the morphologically simple condition; the free stem words with semantically opaque prefixes were numerically faster but not statistically different from the simple. They were also not statistically different from the semantically transparent free stem condition. I interpreted these results as evidence that the relationships between complex words and their component morphemes can be gradient; the more linguistic evidence available to support a complex analysis of a word, the stronger the appearance of morphological effects in behavioral tasks.

I also interpreted these results as evidence that the segment shifting task forces participants to attend to the morphological units of a whole word. When a portion of the source word is highlighted, a visual structure is imposed on the word. This visual structure, which is coextensive with a (pseudo)affix, suggests that the word may contain morphological structure. Participants must evaluate whether the visual structure corresponds to any real lexical structure before they can shift the string. When the word provides clear linguistic evidence for complexity, the string can quickly be identified as a morpheme and shifting is then facilitated. When a word provides only moderate evidence for complexity due to partial semantic opacity, it takes longer for the participant to identify the string as a morpheme; thus, shifting is slightly delayed.

8.2.2 Bound root words

Bound root words are particularly interesting because they tend to lack consistent semantics across forms. Speakers often fail to note a relationship between words that share a bound root, and many prior experimental investigations have failed to find evidence that they include any morphological structure in their lexical representations (e.g., Marslen-Wilson et al. 1994). I also found no evidence of morphological complexity when I investigated bound root words as a homogeneous set. However, I also investigated two additional characteristics of bound root words, namely the semantic transparency of the prefix and the morphophonological behavior of the root. The former proved to be an interacting factor with complexity.

8.2.2.1 The role of form

Experiment 3 revealed a significant difference between bound root words that alternate and bound root words that do not; however, the direction of the effect did not support the hypothesis that roots that alternate are easier to identify as morphemes than roots that do not. Instead, words with roots that maintained a constant phonological form across all their instances shifted significantly faster than words with roots that alternated. By hypothesis, if a root morpheme alternates in form across two instances, there will be less evidence that the two words are related. Consequently, the morphological links across words will be stronger when the roots' form is constant across all instances.

Although having a constant phonological form speeds segment shifting relative to when the form is mutable, this factor was not sufficient to make the bound root words in either condition statistically different from the morphologically simple control condition. Thus, I proposed that phonological constancy was not a strong source of evidence for morphemehood.

8.2.2.2 The role of meaning

151

As with free stem words, I revealed a strong influence of the semantic transparency of a prefix on shifting times for bound root words. While bound roots tend not to provide consistent semantics, I argued that prefixes on bound root words often do. I found that bound root words with semantically transparent prefixes elicited longer shifting times than the morphologically simple control condition. The bound root words with semantically opaque prefixes were no different from morphologically simple words. I interpreted the longer shifting times in the transparent prefix condition as evidence that bound root words can be represented as complex if sufficient linguistic evidence is available to support lexical relationships between a prefixed word and the prefix's representation.

While free stem words with transparent prefixes shifted faster than the simple control words, the bound root words with transparent prefixes shifted slower than the simple controls. I argued that this pattern of results supports a complex analysis for bound root words despite the fact that the bound root words did not pattern the same as the free stem words. I argued that prefix shifting is facilitated when the participant can easily identify the highlighted unit as a morpheme. The process of identifying a unit as a morpheme is conducted by comparing the structure visually imposed by the task to the structure maintained in the lexical representation of the word. When the word provides weak evidence or inconsistent evidence for complexity then that evaluation process is slowed and, in turn, so are the segment shifting times. In contrast, when the word provides consistently strong evidence for or against complexity, then the evaluation process is speeded. Then, the separability of the unit, whether or not the unit is morphemic, becomes relevant.

8.2.2.3 Sensitivity of the segment shifting task

Prior investigations into the representations of morphological information come primarily from two sources: priming studies and nonce word studies. Priming studies are problematic

because it is difficult to disambiguate the effects of form and meaning overlap from the independent contribution of morphological information. Nonce word studies are problematic because the processing of nonce words may be completely distinct from the processing of real words. Nonce words have no lexical representations; thus it is difficult to analogize from a model of the representations and processing of nonce words to a model of the representations and processing of real complex words. The present studies avoid the shortcomings of these other approaches. The segment shifting task was designed to be sensitive to the internal structure of morphologically complex and simple words. The source and response words are not phonologically or semantically related, so the differences in response times are not due to semantic or phonological priming effects. Furthermore, since both source and response words are real words, the results are easily extendable to the (idealized) lexicon of English, rather than being restricted to nonce words.

The present studies have also supplied additional insight into how the segment shifting task works. First, while Feldman and her colleagues argue that the segment shifting task is insensitive to frequency differences across conditions, I found evidence that frequency did affect the task's ability to discriminate between morphological units from non-morphological units. When the frequency of source words was mismatched, the segment shifting task failed to replicate the basic finding that real morphological units are more separable than phonologically matched non-morphemic units (cf., Experiment 3). Most of the words used in my experiments were relatively low frequency, i.e., between 10 and 30 occurrences per million. Feldman's English materials included fairly high frequency words like *city* and *garden.* As a result, equivalent sized frequency differences in her studies may have produced less of an effect, since frequency effects operate on a logarithmic scale. It is also possible that frequency effects are

greater when a task focuses on word beginnings rather than when the task focuses on word endings (Tyler 1984).

8.2.2.4 The representations of prefixes

Although this dissertation argues for the inclusion of morphological information in lexical representations, it also supports the view that semantics is of central importance to morphology. Words consisting of semantically opaque bound roots, as well as cranberry morphemes were, in large part, the motivation for the redefinition of the morpheme and the subsequent separation of the criterial characteristics of the morpheme, namely semantics and phonological form (cf., Jackendoff 1975, for example). The present research suggests that, at least for prefixes, this separation is flawed. Prefixes, perhaps more than some suffixes, require semantics.

As discussed in section 1.5.1, prefixes in English are different from suffixes in several ways. First, they are only derivational. Second, they do not change the part of speech of their bases, as many suffixes do. Nor is there any other grammatical change associated with prefixes (e.g., changing tense or number). Prefixes also do not trigger idiosyncratic phonological alternations, while some suffixes do. Third, there is no phonological evidence to distinguish a prefix from a pseudoprefix. Rather, the only evidence to support a morphemic analysis of an initial string is the form and semantics that it provides. Thus, semantics is even more critical to the identification of prefixes than it may be for suffixes. This does not mean that languages with more productive or regular morphological systems cannot identify prefixes without semantics. Nor does it mean that English suffixes and roots require semantic evidence to the same degree as prefixes, since other types of evidence, such as grammatical functions or phonological alternations, are often available. Rather, it suggests that for English prefixes, both form and meaning are crucial to their status as morphemes.

8.3 Realizing Morphological Representations

The results from the five segment shifting studies support a three-way contrast for how morphological structure is represented in the lexicon. Free stem words with transparent prefixes are fully decomposed, with both morphemes strongly associated with their component parts. Free stem words with opaque prefixes and bound root words with transparent prefixes are both represented as complex as well, but only a portion of the word is strongly associated with its morpheme representation; the remainder of the word is unassociated, or weakly associated, to its morpheme representation. Bound root words with opaque prefixes may be representationally indistinct from morphologically simple words with neither of the morphological components in the word being associated to a sublexical representation. Alternatively, a bound root word with opaque prefixes may have connections to its sublexical components, but all connections are weak. Weak connections may not be sufficient to distinguish a bound root word with an opaque prefix from a morphologically simple word, especially for a task like the segment shifting task.

The speech error elicitation study, however, did not support this three-way distinction. Rather, it provided no evidence for a contrast between the way complexity is represented for bound root words and free stem words; both conditions produced equivalent numbers and types of errors. Furthermore, the semantic relatedness of the presentation words to their intruder words was not a factor in predicting when errors would occur. However, this semantic relatedness study, which rated the semantic transparency across words, was not analogous to the semantic transparency rating study conducted for Experiments 4 and 5, which rated semantic transparency within words. Whether the semantic transparency within a word was a factor for the slips elicitation task is unclear. In sum, the speech error elicitation task, which imposes fewer artificial demands on the participant, did not support the 3-way contrast. Rather, it suggests that the differences in shifting times between bound root and free stem words as well as
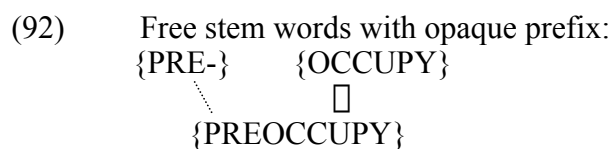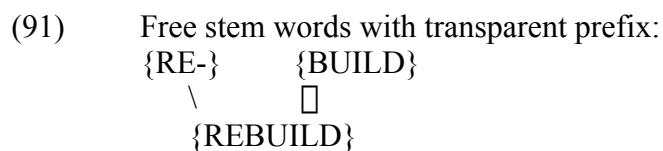
between words with transparent and opaque prefixes were due to evaluative demands imposed by the segment shifting task.

If semantically opaque morphemes contained no relationship to their morpheme representations then bound root words with opaque prefixes should be representationally indistinct from the morphologically simple condition. Since the results from the speech error elicitation task did not reveal a difference between complex words derived from free stems and complex words derived from bound roots, they do not support the proposal that bound root words with opaque prefixes are representationally indistinct from morphologically simple words. Thus, rather than suggest that some words are unassociated to their morphological subparts, the slips data support lexical relationships between complex words and the representations for their component morphemes that are gradient; some morphemes are more strongly associated with the words in which they occur than with others. The segment shifting task is less sensitive to the weaker relationships than the slips elicitation task. The segment shifting task forces participants to attend to units; the visual structure imposed by the task onto the source word must be compared to the lexical representation for the word. Participants must evaluate whether the two structures, the imposed structure and the actual lexical structure, match. Free stem words with transparent prefixes provide clear evidence for complexity and matching the structures occurs quickly. Free stem words with opaque prefixes provide slightly less evidence, so matching structures occurs less quickly. Bound root words with transparent prefixes provide some evidence for complexity, but not from the root, which is the core of the word. Thus, more time is required to determine whether the highlighted unit corresponds to a morpheme or not. Bound root words with opaque prefixes contain only weak associations between the word components and the morpheme representations. The segment shifting task is less sensitive to weak associations, which results in slightly slower shifting times. Morphologically simple words

provide no associations to component morphemes. Therefore, participants can quickly determine that the string is non-morphemic. However, since the string is non-morphemic, its shifting time is slower than the morphemic strings from the free stem conditions.

The slips elicitation task does not force the participants to attend to morphological structure; they merely name the words they read on the screen. Thus, the strength and consistency of evidence for complexity do not arise as crucial factors in this task. Rather, sublexical units are activated via the connections between words and their morphological components. Even weak connections serve to activate the sublexical units sufficiently for them to participate in morphological slips.

The model of morphological representations developed to account for the pattern of data presented in Experiments 1-6 was based on the notion of <u>evidence</u> in support of a morphemic analysis for each component of a word. The present hypothesis suggests that morphological relationships between words can be gradient. Weaker links provide less evidence for complexity than stronger links. Examples (91)-(95) provide schematic representations of the different types of morphologically complex and simple words investigated in this dissertation. These schemas are intended to be fairly atheoretical; the affixes depicted in the diagrams can correspond to explicit affix representations or to Word Formation Rules which assign the morphs that occur for each word.

(91)   Free stem words with transparent prefix:
         {RE-}   {BUILD}
             \   /
            {REBUILD}

(92)   Free stem words with opaque prefix:
         {PRE-}   {OCCUPY}
              ⋮   /
            {PREOCCUPY}

(93)   Bound root words with transparent prefix:
         {RE-}   {-CEDE}
             \   ╱───────────
         {RECEDE}   ⋯   {CONCEDE}

(94)   Bound root words with opaque prefix:
         {DE-}   {-CEIVE}
             ⋮   ╱───────────
         {DECEIVE}   ⋯   {RECEIVE}

(95)   Morphologically simple words with pseudo-prefix:
         {RE-}

         {REMORSE}

Example (91) supplies a schema for how free stem words with transparent prefixes are

represented. The complex word is composed of two morphemic components. Each of these

components has both formal and semantic evidence to suggest it is a morpheme. Thus, the

prefix is strongly associated (schematically depicted with a solid line) to its independent

representation or to the WFR that derived the word. The stem is strongly associated to the

representation for the free standing base. Contrast (91) with (92). The word in (92) has the

same evidence to associate the stem to its free standing base, but the prefix is lacking substantial

semantic content. Thus, while the stem is strongly linked to its base, the prefix on the word and

the prefix representation are only weakly associated (depicted with a dotted line). Example (93)

has the reverse pattern. The prefix, which is transparent, is strongly associated to its prefix

representation. But the root, which is semantically opaque, is weakly associated to its

independent morpheme representation and weakly associated to other words which share the

root unless a common semantic component is evident.[30] The representation for (94), which

consists of two opaque morphemes, has only weak associations to related words and sublexical

---

[30] I remain agnostic as to whether there are independent root representations in the lexicon. When two bound root words share a semantic component, such as *increase* and *decrease,* they may be associated to each other or they may each be associated to the independent root representation, or both. I have no evidence that pertains to this issue.

units. In contrast, the representation in (95) has no evidence to relate the components or pseudocomponents to the morpheme representations. Thus, this word is represented as morphologically simple. It is also possible that bound root words with a semantically opaque prefix might have so little evidence to support a connection between the word and its morpheme representations that it is represented like the simple word of (95).

The representations proposed above have the following characteristics:

Permanent memory representations are maintained for complex words; they are not derived on-line for each usage.
Affixes have some form of representation in the lexicon.
Links are imposed between complex words and their component morphemes, related words, and bases if sufficient evidence for the relation exists.
Links are gradient.

Additionally,

The proposed representations make no claims about whether bound roots have lexical representations independent of the words in which they occur.
Affixes may be separate autonomous lexemes or assigned by a WFR.
The links in this model may be similar to links in IM models in that activation can be spread from representation to representation across the links.

The proposed representations are consistent with the general assumptions of both Item & Process and Item & Arrangement models of morphology since no crucial claim rests on what the nature of the affix representations are or on whether bound roots have independent representations. However, the proposal does assume that complex words have memory representations in the lexicon, which is inconsistent with many specific IA models (e.g., Lieber 1980, 1993). If prefixed words were derived on-line rather than stored as complex representations then there would be no way to explain a difference in shifting times based on stem type or prefix transparency. Thus, this model assumes that complex words have lexical representations, whereas Lieber's model does not. It is possible that the segment shifting and the WOC slips elicitation tasks are sensitive to the structure of the output of morphology, as

suggested in section 1.3, not to the memory representation. But given that the source words are not technically outputted, this explanation seems unlikely.

The proposed representations are also consistent with the main tenets of both morpheme-based and word-based models of the lexicon since my model makes no crucial claims about what types of lexical representations and WFRs are allowed. The model is consistent with claims that only words have lexical representations or that words and morphemes have lexical representations. However, if a word-based view is adopted, the affixes must be understood as rules associated with the representations of lexical items. My model explains the differential effects of shifting times found across varying types of prefixed words via links between words and their components. Thus, the links, or associations, are an important part of what we know about a word and how that word will be treated in behavioral tasks. The proposed representations are clearly consistent with the generative rules proposed by Aronoff (1976) and can be consistent with the redundancy rules proposed by Jackendoff (1975) so long as the rules are linked to the lexical items which they serve to analyze.

Furthermore, Bybee's (1988, 1995) representational model of the lexicon also fails to account for the differences in shifting times across conditions because it does not include a model of lexical processing. Bybee's model has no morphological units, only connections between common linguistic features. Since Bybee's models is of a static lexicon, no *units* are discernible. While she may be able to model acquisition data and intuitions about word relatedness, her model cannot currently account for shifting time differences in bound root and free stem words with transparent prefixes since units and boundaries have no status in her model. Her model would need to include relational information for components within a word in order to capture this difference.

Thus, while my data do not distinguish between families of theories, they do serve to restrict the set of possible models that are compatible with them. While the three models presented above encounter problems when trying to explain the data patterns from Experiments 1-6, many other theories are fully consistent these data. Furthermore, many of the inconsistent theories could account for my data with minor modifications, such as the inclusion of permanent memory representations for complex words.

Though my model may look different from the theories presented in section 1.3, it is not completely new. For example, Aronoff (1976) proposed that word formation is word- based; bound roots do not have independent lexical representations. That is consistent with the proposed model, although I do not take a stand on the issue. Furthermore, he also proposes that contrasts between a compositional meaning of a word and a lexicalized meaning of the word could be reflected in a lexical representation by varying the amount of internal morphological information included in the representation. That is also consistent with my proposal. Recall example (27) from section 1.3, repeated here in example (96).

(96) a. $[[\text{pro=hibit+}]_v +\text{ion}]_n$ "The act of prohibiting something"
     b. $[\text{pro=hibit+tion}]_n$ "The period in the 1920's when alcohol was prohibited"

This example essentially represents the same type of contrast that I am trying to express with the representations in (91)-(95). Crucially, the difference between (96a) and (96b) is in the degree of independence each morpheme has from the word. This is the same type of contrast expressed in (91)-(95). Each morpheme is relatively more independent in example (96a) than the corresponding morphemes in (96b). For Aronoff, this contrast in the degree of independence is represented by including the derivational history for a word. What the derivational history actually corresponds to is inclusion of a link back to the Word Formation Rule that derived or analyzed that portion of the word.

In the notational system proposed above, this contrast would be expressed with the following representations:

(97)   {PROHIBIT} {-ION}
         \\   /
       {PROHIBITION$_1$} — {PROHIBITION$_2$}

The compositional version of *prohibition* would be lexically related to its base and suffix. (The base, *prohibit*, may also have lexical relations to its prefix {pro-} and the root {- hibit} or to other words with the same root, such as *inhibit*.) The non-compositional version of *prohibition,* in contrast, is related only to its compositional counterpart, not directly to its component morphemes.

As mentioned in section 3.2.1, Feldman and her colleagues generally discuss their results in terms of an EM model. The present discussion has followed that tradition. However, I should be clear that the results are also compatible with IM models.

In section 1.5.3, I proposed that an examination of the prefix shifting times for bound root words would be useful for discriminating between IM and EM models. The failure to find a difference between the bound root condition and the morphologically simple condition initially supported an IM model. Furthermore, the fact that roots with constant, non-alternating, phonological forms elicited faster segment shifting times relative to roots that participated in phonological alternations was also predicted by the IM models.

Gradient effects are very common in most IM models. Representations at different levels and within levels are linked via weighted connections. As a result, graded effects are often argued for in place of strict dichotomies.   It is possible, however, that a model of the lexicon that lacks an explicit contrast between free stems and bound roots would have difficulty accounting for the interaction of complexity with transparency found in Experiments 4 and 5.

8.4 Further Issues

As introduced in sections 1.4 and 1.5, one of the most controversial debates in the morphological literature is whether there is any need for the explicit inclusion of morphological information in lexical representations. Since many morphemes are adequately defined as 'a smallest unit of form and meaning' it has been argued by Implied Morpheme models that morphological information is epiphenomenal. Each lexeme must include information about its form and its meaning. These two pieces of information are arbitrary. However, morphological information is derivable from the form and meaning information. IM models argue that the ability to create and understand new words can also be achieved without morphemes. Thus, IM models argue that morphemes are unnecessary and that the appearance of morphological effects are actually due to the combined role of semantic and form information. There are no morphemes in the IM model of the lexicon, only words. In contrast, Explicit Morpheme models argue that much more predictable information can be removed from the lexicon if morphological information is explicitly included. More importantly, morphemes allow for the creation and understanding of novel words. Finally, morphemes are not limited to units of form and meaning.

One of the aims of this dissertation was to provide some new evidence to add to this debate. Unfortunately, none of my results seems to clearly distinguish between these two models. I do find evidence for the central importance of semantics in identifying morphological complexity in prefixed words, which is a claim maintained by the IM models. In addition, I did not find evidence that a phonological alternation was a crucial characteristic to identifying a bound root as a morpheme, a claim maintained by many EM theorists. While the distinct behavior of transparent and opaque prefixes on free stem words and bound root words does seem to suggest the need for words to be associated to morpheme representations, many IM models can also

account for this data with the correct architecture for their network of nodes and excitatory and inhibitory connections.

8.4.1 Bracketed structure

Many linguistic theories posit that, as a word progresses through successive cycles of word formation, information about previous derivations is lost (e.g., Kiparsky 1982, Aronoff 1976). For example, the internal linguistic structure for a word like *unbelievable* might not be *un-believe-able*, with a flat structure, but *un-believable[31]* or *(un+(believe+able))*. One way of implementing the loss of derivational history is with a process called *bracket erasure*. Part of the rationale for this mechanism is that affixes generally do not need to refer to any word properties internal to the most recent derivation. In other words, no selectional restriction of an affix refers to past properties of the stem it wants to combine with. For example, the suffix *-ly,* which only attaches to adjectives, never interrogates a word's derivational history to see if it was ever an adjective, it only cares about the current state of the word. The word a*ngrily* can be formed because *angry* is an adjective regardless of the fact that the stem *anger* is a noun. In contrast, *quickenly* cannot be formed even though the stem *quick* is an adjective because the outermost word, *quicken*, is a verb. The claims of bracket erasure suggest that not all morphological boundaries are equivalent; the boundary between *believe* and *able* in *believable* is much more prominent than in the word *unbelievable*. In terms of the notational system presented above, these two words would be represented as in (98).

(98)   {UN-}   {BELIEVE} {-ABLE}
              \   \   /
          {UNBELIEVABLE}   —   {BELIEVABLE}

---

[31] The structure *unbeliev-able* is ruled out because *unbelieve* is not a word. *Believe* is a verb and the prefix *un-* attaches to adjectives. Thus, the adjective-deriving suffix *-able* must first be added, creating the adjective *believable*.

The diagram in (98) illustrates that the suffix {-able} is directly associated only to the bimorphemic word *believable* but not directly related to the polymorphemic word *unblievable.* The stem {believe} is also only directly related to the bimorphemic word *believable.* The polymorphemic word *unbelievable* is associated to the prefix {un-} and the bimorphemic base {beliveable}; it is only indirectly related to the suffix {-able}.

 The shifting times for suffixes like {-able} could be compared when they occurred in source words like *unbelievable* vs. *believable.* Since the suffix is internal to another word formation process in *unbelievable*, I hypothesize that it should produce slower shifting times than the same suffix in a source word like *believable*, in which the suffix is not contained within a later derivational process. I will refer to suffixes on words like *unbelievable* as 'internal' and suffixes on words like *believable* as 'external'. Both of these two complex conditions would be contrasted to a morphologically simple control condition with pseudosuffixes, such as *turntable*, although two control conditions might be needed to account for the string length confound inherent to the internal/external morpheme contrast. If this contrast reveals a reliable difference in the shifting times for internal and external suffixes, it will provide strong evidence for morphological information in the lexicon. Because there are no morphemes in IM models, morphological effects should be based on the degree of form and semantic overlap. Since the semantic and formal evidence for the morphemes {believe} and {-able} in *unbelievable* and *believable* are the same, IM models would have difficulty explaining a difference between internal and external suffix shifting times.

8.4.2 Contribution of the semantics of bound roots

Another issue that was not investigated with the segment shifting task is whether the semantic contribution of the bound root plays any role in the shifting times for prefixes. While most bound root words are semantically opaque and the words that they occur in are relatively

unrelated, some roots do seem to provide consistent semantics. For example, the words *receive and deceive* are semantically unrelated but the words *recede and proceed* have a common semantic component, although it might be hard to articulate. Even more related are opposites like *increase* and *decrease.* While I investigated the contribution of prefix semantics, I generally ignored the role of root semantics. Following the hypothesis proposed in section 6.3.3, semantically transparent bound roots should be strongly associated to the lexical representation of the root or to other words that share the root, as illustrated in the schematic representation provided in (99). Contrast the strong lexical relationship represented in example (99) to the weak associations in example (94) above. If a word has both a transparent prefix and a transparent root, it is predicted to shift faster than the simple controls. These words, which have no conflicting cues for complexity, should be equally as fast as the free stem words with transparent prefixes. While this manipulation has not yet been conducted, its relevance to the current model is clear.

(99)   Bound root words whose root and prefix are transparent:
         {IN-}
           |
         {INCREASE} — {DECREASE}

8.5 Conclusion

To summarize, this dissertation has provided evidence for the inclusion of independent morpheme representations associated with prefixes in free stem words as well as bound root words with transparent prefixes. I have introduced the notion of prefix transparency as a contributing factor for the morphological analysis of words in the lexicon. Furthermore, I suggested that phonological alternations are not diagnostics to identify a bound root as a morpheme when semantics is absent, contrary to Aronoff's proposal (1976). I presented a model of morphological complexity in which complex words are associated to their component

morpheme representations via weighted links. The effect of semantic transparency of prefixes in the segment shifting task was argued to be the result of conflicting cues for complexity due to weak morphological connections in the lexicon. While the results presented in this dissertation do not answer all the questions presented in the introduction, they do restrict the way morphological information can be represented in the lexicon as well as what the criteria for morphological complexity are in English.

Appendix A

Source words for Experiment 1

SIMPLE	COMPLEX
HUNGER	SINGER
ARMY	JUMPY
BANISH	BOYISH
MESSAGE	BAGGAGE
FIGMENT	AILMENT
VENTURE	MOISTURE
BUTTER	BATTER
TIDY	FROSTY
SANDAL	RENTAL
SEWAGE	COVERAGE
PUNISH	GIRLISH
FOREST	HIGHEST
CASTOR	SCULPTOR
PIGMENT	ADJUSTMENT
PITY	LACY
QUARTER	STARTER
INNING	WINNING
MANAGE	DRAINAGE
PASTURE	FAILURE
LOBSTER	MOBSTER
RADISH	FOOLISH
MITTEN	BITTEN
PARCHMENT	PAVEMENT

POWER	MOWER
CANNY	SHADY
EARNEST	LARGEST
FINISH	CHILDISH
FENDER	STOPPER
GARBAGE	DOSAGE
DOZEN	DRUNKEN
MYSTERY	BAKERY
VANISH	SLUGGISH
BELLY	FOGGY
CENTER	COLDER

Appendix B

Source Words for Experiment 2

BOUND	FREE	SIMPLE
RECEDE	RETRACE	REMORSE
DEDUCE	DEFORMED	DELIVER
INCLUDE	INADEQUATE	INJURY
PROCEED	PROLONG	PROPRIETY
EXCESS	EXHUSBAND	EXERCISE
SUBSCRIBE	SUBGROUP	SUBLIME
IMPART	IMMORAL	IMITATE
COHESIVE	COOPERATE	COCONUT
PRECEDE	PREPAID	PREMIER
ENTIRE	ENJOY	ENERGY
REFER	REBORN	REPAIR
DISDAIN	DISCONTENT	DISTRIBUTE
EMPATHY	EMBODIED	EMPIRE
ILLUSIONS	ILLEGAL	ILLUSTRATE
INGEST	INATTENTIVE	INNOCENT
DECEIVE	DEFROST	DENIAL
PROFANE	PROCHOICE	PROLIFIC
RESULT	REACTION	RELIGION
IMPOSE	IMPARTIAL	IMPUDENT
REFLECT	REBUILD	RELAX
SUBJECT	SUBDIVISION	SUBSTITUTE
INSPECT	INDEBTED	INDIANS
PREDICT	PREHISTORIC	PREACHER
DECLINE	DEMORALIZE	DELIGHT

Appendix C

Source Words for Experiment 3

NON-ALTERNATING	ALTERNATING	FREE	SIMPLE
DETACH	DECLINE	DECOMPOSE	DELIGHT
EXPLORE	EXPLODE	EXCLAIM	EXAMINE
DISCERN	DISRUPT	DISAPPROVE	DISGUST
INSULT	INCLUDE	INADEQUATE	INJURY
PRESERVE	PRECEDE	PREPAID	PREMIER
RETRIEVE	REFLECT	REFUEL	RELAX
SUBSIST	SUBSCRIBE	SUBGROUP	SUBPOENA
PROPEL	PROCEED	PROCHOICE	PROLIFIC
INVOLVE	INVENT	INSIDE	INDIAN
DEPEND	DEDUCE	DEFROST	DENIAL
IMPACT	IMPART	IMMORAL	IMITATE
REFUSE	REFER	REBORN	REPAIR
DISDAIN	DISCRETE	DISARM	DISASTER

Appendix D

Source Words for Experiment 4

TRANSPARENT-FREE	OPAQUE - FREE	SIMPLE
DISBAR	DISBELIEVE	DISGUST
IMMORTAL	IMMATURE	IMPUDENT
REBUILD	REACT	REGAL
PROCHOICE	PROCLAIM	PROLIFIC
ENABLE	ENSLAVE	ENTAIL
DEMORALIZE	DEHYDRATE	DECLARE
DISAPPROVE	DISFIGURE	DISCIPLE
MISCOUNT	MISCARRY	MISCHIEF
EXCLAIM	EXCHANGE	EXERT
ENCIRCLE	ENDEAR	ENAMEL
PRECAUTION	PREOCCUPY	PREACHER

Appendix E

Source Words for Experiment 5

TRANSPARENT- BOUND	OPAQUE-BOUND	SIMPLE
REFLECT	RESPECT	REMORSE
PROVIDE	PROHIBIT	PROLIFIC
PRECEDE	PRETEND	PREMIER
INJECT	INSULT	INVITE
DECLINE	DECEIVE	DENIAL
EXPLODE	EXPLORE	EXERCISE
DISPOSE	DISCRETE	DISGUST
IMPACT	IMPART	IMPUDENT
PREDICT	PRESENT	PREACHER
RECEDE	REFUSE	RELATE
DISPEL	DISCERN	DISASTER
SUBMIT	SUBSCRIBE	SUBPOENA
PERSIST	PERMIT	PERSON

Appendix F

Presentation word pairs for Experiment 6

FREE	BOUND	SIMPLE
Unused misnamed	Deflect recline	Widow shallow
Inaction transform	Project succeed	Remedy colic
Pretrial misjudged	Reject induce	Mugger baffled
Untitled subconscious	Remit advise	Sector fashion
Prepayment coexist	Reclusive insistent	Matron pallet
Rewind undone	Perturb dissuade	Legion regal
Unfold ensure	Disruption contortion	Rattle morbid
Retest preview	Dispose propel	Shower flatter
Reappear discharge	Export implode	Money hunky
Misshaped realign	Expose oppress	Certify forgery
Recite exclaim	Inflate decrease	Ditches sweater

References

Anshen, F., & Aronoff, M. (1988). Producing morphologically complex words. *Linguistics*, *26*, 641-655.

Aronoff, M. (1976). *Word formation in generative grammar*. Cambridge, Mass: MIT Press.

Baars, B., & MacKay, D. (1978). Experimentally eliciting phonetic and sentential speech errors: Methods, implications, and work in progress. *Language in Society*, *7*, 105-109.

Baars, B., Mattson, M., & Cruickshank, G. (1985). *The induction of syllable exchanges*. Unpublished manuscript, Department of Psychology, State University of New York, Stony Brook.

Baars, B., & Motley, M. (1976). Spoonerisms as sequencer conflicts: Evidence from artificially elicited errors. *American Journal of Psychology*, *89*, 467-484.

Baars, B. (1977). The planning of speech: Is there semantic editing prior to speech articulation? [diss], Los Angeles, CA: UCLA.

Baars, B. (1992). A dozen competing-Plans techniques for inducing predictable slips in speech and action. In B. Baars (Ed.), *Experimental slips and human error: Exploring the architecture of volition* (pp. 129-150). New York: Plenum Press.

Baars, B., Motley, M., & MacKay, D. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*, 382-391.

Baayen, R. H. (1994). Productivity in language production. *Language & Cognitive Processes: Morphological structure, lexical representation and lexical access*, *9*, 447-469.

Beauvillain, C. (1994). Morphological structure in visual word recognition: Evidence from prefixed and suffixed words. *Language & Cognitive Processes: Morphological structure, lexical representation and lexical access*, *9*, 317-339.

Beauvillain, C., & Segui, J. (1992). Representation and processing of morphological information. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning. Advances in psychology,* (pp. 377-388). Amsterdam, Netherlands: North-Holland.

Bentin, S. (1989). Orthography and phonology in lexical decision: Evidence from repetition effects at different lags. *Journal of Experimental psychology: Learning, Memory & Cognition*, *15*, 61-72.

Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags. *The quarterly journal of experimental psychology*, *42*, 693-711.

Bertram, R., Schreuder, R., & Baayen, R. H. (2000a). The balance of storage and computation in morphological processing: The role of word formation. *Journal of Experimental psychology: Learning, Memory & Cognition*, *26*, 489-511.

Bertram, R., Baayen, R. H., & Schreuder, R. (2000b). Effects of family size for complex words. *Journal of Memory & Language*, *42*, 390-405.

Bloomfield, L. (1933). *Language*. New York: Holt.

Bochner, H. (1993). *Simplicity in Generative Morphology*. Publications in Language Sciences, vol. 37. New York: Mouton de Gruyter.

Balota, D., & Chumbley. (1985). The locus of word-frequency effects in the pronunciation task. *Journal of Memory & Language*, *24*, 89-106.

Bybee, J. (1988). Morphology as Lexical Organization. In M. N. Hammond, M. (Ed.), *Theoretical Morphology*. Academic Press.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*, 425-255.

Bybee, J., & Slobin, D. (1982). Rules and schemas in the development and use of English past tense. *Language*, *59*, 251-270.

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

Caramazza, A., Miceli, G., Silveri, M. C., & Laudanna, A. (1985). Reading mechanisms and the organisation of the lexicon: Evidence from. *Cognitive Neuropsychology*, *21*, 81-114.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper and Row.

Connine, C., Blasko, D., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory & Language*, *32*, 193-210.

Cutler, A. (1980). Errors of stress and intonation. In V. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 67-80). New York: Academic Press.

Cutler, A., & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 133-142.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 113-121.

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

Denning, K. & Leben, W. (1995). *English Vocabulary Elements*, Oxford:Oxford University Press.

Drews, E., & Zwitserlood, P. (1995). Morphological and orthographic similarity in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1098-1116.

Emmorey, K. (1989). Auditory morphological priming in the lexicon. *Language and Cognitive Processes*, *4*(2), 73-92.

Feldman L. B., & Fowler, C. (1987). *Morphemic segments shift faster than nonmorphemic controls*. Paper. Psychonomic Society, Seattle, WA.

Feldman, L. B. (1991). The contribution of morphology to word recognition. *Psychological Research*, *51*, 33-41.

Feldman, L. B., Frost, R., & Pnini, T. (1995). Decomposing words into their constituent morphemes: Evidence from English and Hebrew. *Journal of Experimental psychology: Learning, Memory & Cognition*, *21*, 947-960.

Feldman, L. B., & Soltano, E. (1999). Morphological priming: The role of prime duration, semantic transparency, and affix position. *Brain & Language*, *68*, 33-39.

Forster, K., & Chambers, S. (1973). *Journal of Verbal Learning and Verbal Behavior*, *12*, 627-635.

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

Fromkin, V. (1971). The nonanomalous nature of anomalous utterances. *Language*, *47*, 27-52.

Fromkin, V. (1973). *Speech errors as linguistic evidence*. The Hague: Mouton.

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological. *Journal of Experimental psychology: Learning, Memory & Cognition*, *23*, 829-856.

Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production*. London: Academic Press.

Goldinger, S., Luce, P., & Pisoni, D. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory & Language*, *28*, 501-518.

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

Grossberg, S., Boardman, I., & Cohen, M. (1997). Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 483-503.

Hall, C. J. (1995). Formal linguistics and mental representation: Psycholinguistic contributions to the identification and explanation of morphological and syntactic competence. *Language and Cognitive Processes*, *10*, 169-187.

Halle, M. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry*, *4*, 3-16.

Hockett, C. (1954). Two models of grammatical description. *Word*, *10*, 210-31.

Inhoff, A. W., Briihl, D., & Schwartz, J. (1996). Compound word effects differ in reading, on-line naming and delayed naming tasks. *Memory & Cognition*, *24*, 466-476.

Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language*, *51*, 637-71.

Jenson, J. (1990). *Morphology: Word structure in generative grammar*. Amsterdam: John Benjamins.

Kempley, S., & Morton, J. (1982). The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology*, *13*, 441-454.

Kiparsky, P. (1982). Lexical phonology and morphology. In I. S. Yang (Ed.), *Linguistics in the morning calm* (pp. 3-91). Seoul: Hanshin.

Kowamoto, A. H. (1988). Distributed representations of ambiguous words and their resolution in a connectionist network. In S. I. Small, G. W. Cottrell & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution*. San Mateo, CA: Morgan Kaufmann.

Laudanna, A., Badecker, W., & Caramazza, A. (1992). Processing inflectional and derivational morphology. *Journal of Memory & Language*, *31*, 333-348.

Laudanna, A., Cermele, A., & Caramazza, A. (1997). Morpho-lexical Representations in naming. *Language and Cognitive Processes*, *12*, 49-66.

Libben, G. (1993). Are morphological structures computed during word recognition? *Journal of Psycholinguistic Research*, *22*, 535-544.

Lieber, R. (1992). *Deconstructing Morphology: Word formation in syntactic theory*. Chicago: University of Chicago Press.

Lieber, R. (1980). The organization of the lexicon [diss]. MIT.

Lukatela, G., Carello, C., & Turvey, M. (1987). Lexical representation of regular and irregular inflected nouns. *Language and Cognitive Processes*, *2*, 1-17.

Lukatela, G., Gligorijevic, B., Kostic, A., & Turvey, M. (1980). Representation of inflected nouns in the internal lexicon. *Memory & Cognition*, *8*, 415-423.

MacKay, D. (1976). On the retrieval and lexical structure of verbs. *Journal of Verbal Learning and Verbal Behavior*, *15*, 169-182.

MacKay, D. (1978). Derivational rules and the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *17*, 61-71.

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

Marslen-Wilson, W., Zhou, X., & Ford, M. (1996). Morphology, modality, and lexical architecture. In G. E. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1996* (pp. 117-134). Dordrecht: Kluwer Academic Publishers.

McClelland, J., & Rumelhart, D. (1985). Distributed memory and the representation of general and specific knowledge. *Journal of Experimental Psychology: General*, *114*, 159-188.

Meunier, F., & Segui, J. (1999). Frequency effects in auditory word recognition: The case of suffixed words. *Journal of Memory & Language*, *41*, 327-344.

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

Napps, S. (1989). Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory & Cognition*, *17*, 729-739.

Nida, E. (1946). *Morphology: The descriptive analysis of words*. Ann Arbor: University of Michigan Press.

Orsolini, M., & Marslen-Wilson, W. (1997). Universals in morphological representation: Evidence from Italian. *Language and Cognitive Processes*, *12*, 1-47.

Pillion, A. (1998). Morpheme units in speech production: Evidence from laboratory-induced verbal slips. *Language and Cognitive Processes*, *13*, 465-498.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a Parallel Distributed Processing model of language acquisition. *Cognition*, *28*, 73-193.

Plaut, D., McClelland, J. L., Seidenberg, M. S., & Patterson, D. (1996). Understanding normal and impaired word reading. Computational principles in quasi-regular domains. *Psychological Review*, *103*.

Plaut, D., & Gonnerman, L. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes*, 1-19.

Plaut, D., & McClelland, J. (1993). Generalization with componential attractors: Word and nonword reading in an attractor network. In *Proceedings of the 15th annual conference of the cognitive science society*. 15th Annual Conference of the Cognitive Science Society. Hillsdale, NJ.: Erlbaum.

Rueckl, J. G., Mikolinski, M., Raveh, M., Miner, C., & Mars, F. (1997). Morphological priming, fragment completion, and connectionist networks. *Journal of Memory & Language*, *36*, 382-405.

Rueckl, J., & Raveh, M. (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain & Language*, *68*, 110-117.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tense of English verbs. In J. McClelland & D. W. Rumelhart (Eds.), *Parallel distributed processing: Volume 2*. Cambridge, MA: MIT Press.

Schreuder, R., & Baayen, R. H. (1995). Modeling Morphological Processing. In L. B. Feldman (Ed.), *Morphological Aspects of Language Processing* (pp. 131-156). Hillsdale, NJ: Lawrence Erlbaum.

Schreuder, R., Grendel, M., Poulisse, N., Roelofs, A., & van de Voort, M. (1990). Lexical processing, morphological complexity and reading. In D. Balota & G. Flores d'Arcais (Eds.), *Comprehension processes in reading* (pp. 125-141). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schriefers, H., Zwitserlood, P., & Roelofs, A. (1991). The identification of morphologically complex words: Continuous processing or decomposition? *Journal of Memory and Language*, *30*, 26-47.

Schriefers, H., Friederici, A., & Graetz, P. (1992). Inflectional and derivational morphology in the mental lexicon: Symmetries and asymmetries in repetition priming. *The Quarterly Journal of Experimental Psychology. Human Experimental Psychology*, *44A*, 373-390.

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

Seidenberg, M. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading* (pp. 245-263). London: Erlbaum.

Selkirk, E. (1982). *The syntax of words*. Cambridge, MA: MIT Press.

Spencer, A. (1991). *Morphological Theory*. Cambridge, Mass: Blackwell Publisher.

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

Stanners, R., Neiser, J., & Painton, S. (1979b). Memory representation for prefixed words. *Journal of verbal learning and verbal behavior*, *18*, 733-743.

Stemberger, J. (1985). Bound morpheme loss errors in normal and agrammatic speech: One mechanism or two? *Brain & Language*, *25*, 246-256.

Stemberger, J., & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition*, *14*, 17-26.

Stolz, J., & Feldman, L. B. (1995). The Role of Orthographic and Semantic Transparency of the Base Morpheme in Morphological Processing. In L. B. Feldman (Ed.), *Morphological Aspects of Language Processing* (pp. 109-129). Hillsdale, NJ: Lawrence Erlbaum.

Stolz, J., & Besner, D. (1998). Levels of representation in visual word recognition: A dissociation between morphological and semantic processing. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1642-1655.

Swinney, K., Onifer, W., Prather, P., & Hirshkowitz, M. (1979). Semantic facilitation across modalities in the processing of individual words and sentences. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645-659.

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. *The Quarterly Journal of Experimental Psychology*, *38*, 351-366.

Taft, M. (1988). A morphological-decompositional model of lexical representations. *Linguistics*, *26*, 657-667.

Tanenhaus, M., Flanigan, H., & Seidenberg, M. (1980). Orthographic and phonological activation in auditory and visual word recognition. *Memory and Cognition*, *8*, 513-520.

Tyler, L. (1984). The structure of the initial cohort: Evidence from gating. *Perception and Psychophysics*, *36*, 217-222.

Tyler, L. K., & Marslen-Wilson, W. (1986). The effects of context on the recognition of polymorphemic words. *Journal of Memory & Language*, *25*, 741-752.

Tyler, L. K., Marslen-Wilson, W., Rentoul, J., & Hanney, P. (1988). Continuous and discontinuous access n spoken word recognition: The role of derivational prefixes. *Journal of Memory & Language*, *27*, 368-381.

Whaley, C. (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, *17*, 143-154.

Wurm, L. H. (1997). Auditory processing of prefixed words is both continuous and decompositional. *Journal of Memory & Language*, *37*, 438-461.

Wurm, L. (2000). Auditory processing of polymorphemic pseudowords. *Journal of Memory & Language*, *42*, 255-271.