Journal of Experimental Psychology: General
2007, Vol. 136, No. 2, 323–345

0096-3445/07/$12.00 DOI: 10.1037/0096-3445.136.2.323

Graded Semantic and Phonological Similarity Effects in Priming: Evidence for a Distributed Connectionist Approach to Morphology

Laura M. Gonnerman
Lehigh University

Mark S. Seidenberg
University of Wisconsin—Madison

Elaine S. Andersen
University of Southern California

A considerable body of empirical and theoretical research suggests that morphological structure governs the representation of words in memory and that many words are decomposed into morphological components in processing. The authors investigated an alternative approach in which morphology arises from the interaction of semantic and phonological codes. A series of cross-modal lexical decision experiments shows that the magnitude of priming reflects the degree of semantic and phonological overlap between words. Crucially, moderately similar items produce intermediate facilitation (e.g., *lately–late*). This pattern is observed for word pairs exhibiting different types of morphological relationships, including suffixed–stem (e.g., *teacher–teach*), suffixed–suffixed (e.g., *saintly–sainthood*), and prefixed–stem pairs (*preheat–heat*). The results can be understood in terms of connectionist models that use distributed representations rather than discrete morphemes.

Keywords: derivational morphology, mental lexicon, cross-modal priming, word recognition, connectionism

One of the fundamental problems in the study of language is to characterize knowledge of words and how this knowledge is used in comprehension and production. The focus of the present article is on *derivational morphology*, the aspect of lexical knowledge concerning the structure and formation of complex words. Words such as *baker* and *talkative* appear to consist of components, traditionally called *morphemes*, that can be recombined to form other words (e.g., *baking*, *talker*). *Baker*, for example, can be analyzed as consisting of the root (or stem) morpheme *bake* and the suffix *–er*. In such words, each of the components contributes to the meaning of the whole. Moreover, *bake* makes similar contributions to related words such as *baked* and *bakery*; the agentive suffix *–er* makes similar contributions to words such as *talker* and *writer*. How this information is acquired, represented, and used has been the focus of considerable research. Most contemporary theories assume that complex words consist of discrete morphemic units that are represented in memory and that are used in processing (see chapters in Feldman, 1995, for some examples). In this article we discuss some of the limitations of this approach and develop an alternative, inspired by connectionist theories of knowledge representation and learning, in which graded, nondiscrete morphological structures emerge in the course of learning relations among the sounds, meanings, and spellings of words. We present four experiments that provide support for this view. The studies involve priming effects that are observed for related words such as *baker–bake* and *lovely–love*. They show that priming effects are a graded function of the degree of semantic and phonological overlap between words rather than an indication of morphological relatedness.

BACKGROUND

Producing and comprehending words involves mappings between form (*phonology* and, in literate individuals, *orthography*) and meaning. For simple words, this mapping is largely arbitrary; the word for “domestic canine” happens to be *dog* but it could as well have been something else, as it is in other languages. In contrast, morphological structure involves nonarbitrary form–meaning correspondences, as shown in the *baker* and *talkative* examples mentioned earlier. Most contemporary theories assume

---

Laura M. Gonnerman, Department of Psychology, Lehigh University; Mark S. Seidenberg, Department of Psychology, University of Wisconsin—Madison; Elaine S. Andersen, Departments of Psychology and Linguistics and Program in Neuroscience, University of Southern California.

This work was financially supported in part by a William Keck Graduate Fellowship, a University of Southern California Dissertation Year Fellowship, and National Institute of Mental Health (NIMH) National Research Service Award F32-DC00374 and National Institutes of Health (NIH) Grant R01-MH55628 to Laura M. Gonnerman. Elaine S. Andersen was supported by NIH Grant R01-AG11773, and Mark S. Seidenberg was supported by NIMH Grant RO1-MH58723 and NIMH Research Scientist Development Award KO2-MH01188. We thank David Plaut, Joseph Devlin, Brian MacWhinney, and the Carnegie Mellon University Parallel Distributed Processing research group for helpful comments and discussion.

Correspondence concerning this article should be addressed to Laura M. Gonnerman, Department of Psychology, Lehigh University, 17 Memorial Drive East, Bethlehem, PA 18015. E-mail: lag5@lehigh.edu

that morphological structure represents a distinct type of linguistic knowledge that encodes information that extends beyond mere correlations between form and meaning. This assumption is based on several observations. First, some words that do not exhibit strong form–meaning correlations nonetheless appear to exhibit internal morphological structure (Aronoff, 1976). For example, *gingerly* appears to pattern with words such as *nicely* and *badly*, but whereas the meanings of *nicely* and *badly* are systematically related to the meanings of their "stem morphemes" (i.e., *nice* and *bad*), the same is not true of *gingerly*. Words such as *gingerly* therefore suggest that morphological structure is not merely a function of the semantic properties of the components. Second, people can understand and produce novel morphologically complex forms even if the stem morpheme is semantically empty. Thus, if you know something can be *zimmed*, you can infer that it is *zimmable*. Finally, a body of experimental research suggests that morphologically structured words produce stronger behavioral effects than do words that are merely semantically or phonologically related. These observations are considered further in the following paragraphs; the main issue is whether they demand the discrete morphemes assumed in previous theories or rather arise from other principles. The plan of the remainder of the article is as follows. We first describe some methodological and conceptual issues that have made morphology a contentious area of research and identify some important limitations of previous studies. We then develop an alternative view and describe four experiments that provide data bearing on the competing accounts. In the General Discussion, we discuss unresolved issues (e.g., whether our approach could extend to languages other than English) and directions for future research.

Psycholinguistic research on morphology has focused on two main issues: the role of morphemes in the storage and processing of words and the even more basic question of what constitutes a morpheme. The major controversy with regard to storage and processing concerns whether words are stored in memory and produced and comprehended as wholes or in terms of component morphemes. The issue about what constitutes a morpheme has centered on the status of unclear cases such as *remit* or *grocer*, which exhibit some but not all properties thought to be characteristic of morphological structure.

Storage and Processing Issues

Three approaches to the processing of morphologically complex words have been proposed: *whole word*, *decompositional*, and *hybrid*. Proponents of the whole word approach (Butterworth, 1983; Manelis & Tharp, 1977) have argued that decomposing complex words is less efficient than processing them as wholes, especially given that decomposition will produce incorrect segmentation in many cases (e.g., a decomposition process that parses *baker* correctly will misanalyze *corner*). This view is difficult to reconcile with the studies we cite in which properties of subword components affected the processing of complex words. The whole word approach also leaves open the question of how novel, complex words are comprehended and produced. Because of these limitations, the whole word approach has largely been abandoned in favor of models that incorporate some form of decomposition.

In an influential study, Taft and Forster (1975) found that lexical decision latencies were longer for pseudoaffixed nonwords containing morphological stems, such as *dejuvenate*, than for pseudoaffixed nonwords such as *depertoire.* The findings were interpreted as evidence for a model in which recognition involves removing affixes in order to recover stem morphemes. In this affix-stripping account, the lexicon is assumed to be organized in terms of the stems that underlie related words; for example, a stem such as *book* provides access to the group of related words in which it occurs. *Juvenate* was treated as a stem morpheme because it occurs in the word *rejuvenate*, which was considered a prefixed word. Processing *dejuvenate* was assumed to involve stripping the prefix *de–*, leaving the stem morpheme *juvenate*, which accesses an entry in the mental lexicon, slowing the decision that *dejuvenate* is not a word. *Pertoire*, in contrast, occurs in *repertoire*, which was assumed to be unprefixed; therefore *pertoire* is not represented in memory and so *depertoire* can be rejected more quickly.

The Taft and Forster (1975) study was followed by many others showing similar effects through the use of a variety of methodologies (see chapters in Feldman, 1995, for overviews). For example, studies in several languages used frequency effects to diagnose the use of morphemic units. The frequency with which a word is used affects how long it takes to comprehend or to produce (e.g., Rayner & Pollatsek, 1989). This effect is standardly interpreted as evidence that words are represented in memory with a record of how often they are used (e.g., Forster & Chambers, 1973). Other studies have shown that latencies to name or to comprehend complex words are also affected by the frequencies of morphemic constituents. In the first study of this type, Taft (1979) manipulated both *surface frequency* (the frequency of a compound word such as *notebook*) and *root morpheme frequency* (the summed frequency of all words assumed to contain a given root, e.g., all words containing *book*). Lexical decision latencies were affected by surface frequency with root frequency equated; however, latencies were also affected by root frequency with surface frequency equated. Burani, Salmaso, and Caramazza (1984) and Meunier and Segui (1999) presented similar results for Italian and French, respectively. Reasoning by analogy to conclusions based on effects of word frequency, researchers have taken effects as evidence that these units are also represented in the mental lexicon and used in lexical access.

On the basis of a large body of such findings, most psycholinguistic theories have incorporated one or another version of the lexical decomposition hypothesis, which makes three main assumptions:

1. Complex words consist of sequences of discrete morphemes.

2. Words are represented in the mental lexicon in terms of component morphemes.

3. Comprehension of many complex words involves decomposing them into these units.

Thus *talker* is represented as *talk* + *er*, and recognition involves procedures for recovering this structure (e.g., suffix stripping). This approach has considerable intuitive appeal. It seems clear from examples such as *baker* and *talker* that words consist of discrete components and from examples such as *geneticize* that people are able to use this knowledge to form new words. Because people possess this knowledge, intuition suggests it is unlikely that they fail to use it in the comprehension or the production of familiar words. Further pretheoretical motivation for this approach was provided by intuitions about how to achieve economy of storage and efficiency of processing (Bradley, 1980; Sandra, 1994). Memory and processing resources were assumed to be limited, and therefore rapid and reliable word recognition could only be achieved by exploiting the redundancies among words. Thus if *baker*, *writer*, and *talker* were all represented in terms of their component morphemes, they could be recognized by simple heuristics such as suffix stripping that decompose words into their parts.

In contrast to morphologically complex words such as *talker* and *nicely*, there are words such as *kangaroo* and *citadel* that do not appear to contain morphological subunits. Insofar as such words cannot be decomposed, they have been thought to require processing mechanisms that operate over whole words. This has led to the development of hybrid models that incorporate both whole word and decompositional mechanisms. Although all hybrid models assume an explicit role for morphological structure, they vary as to which word types are decomposed and whether decomposition applies to storage or processing or both. For example, Andrews (1986) argued that compound words are decomposed in lexical access, suffixed words are optionally decomposed, and all types of complex words are represented as whole forms, whereas Colé et al. (1989) proposed that prefixed words are processed as whole forms and suffixed words are decomposed. In the augmented addressed morphology model of Caramazza and colleagues (Burani & Caramazza, 1987; Caramazza, Laudanna, & Romani, 1988), a whole word procedure applies to monomorphemic words such as *kangaroo* and *tinsel* and a decomposition procedure to complex and novel words. Other hybrid models, such as Frauenfelder and Schreuder's (1992) morphological race model, adopt the further assumption that the whole word and decomposition procedures operate in parallel with a race between them.

In related work, Marslen-Wilson, Tyler, Waksler, and Older (1994) took the important step of proposing that whether words are decomposed or processed holistically depends in part on their semantic properties. Previous research tended to ignore this issue on the view that decomposition procedures were based on the forms of words rather than on their meanings. Marslen-Wilson et al. proposed that semantically transparent morphologically complex words such as *government* are represented and accessed in terms of their components, whereas words such as *department*, which are formally similar but semantically opaque, are represented as wholes. A series of cross-modal priming studies provided evidence consistent with this account; these studies, which are discussed in detail in the following section, provided the basis for our experiments.

The Problem With Morphemes

Our research addresses the central problem for theories that incorporate lexical decomposition: They assume that complex words consist of discrete morphemes, but there is little agreement about how morphemes are defined. The concept of discrete morphemes seems adequate as long as attention is focused on clear cases that emphasize the contrast between morphologically simple and complex words, such as *kangaroo* and *baker*, respectively. As we have noted, however, languages such as English admit many words that exhibit partial regularities. Considering the broader range of cases leads to a different conception of the basis of morphological structure and its role in processing.

There is little agreement about the criteria that determine a word's morphological status. In structural linguistic theories, morphemes are defined as minimal meaning-bearing units (Hockett, 1958) that are arrayed like beads on a string. The meaning of a complex word is therefore predictable from the meanings of its component morphemes. Many psycholinguistic models have accepted this characterization, making the further assumption that morphological knowledge is acquired in the course of language development, encoded in memory, and used in processing. However, problems with the classical concept of morpheme have been widely recognized since the work of Aronoff (1976). The basic issue is that many words deviate from the beads-on-a-string characterization: They exhibit some but not all of the properties assumed under this characterization. Consider the following examples, which illustrate patterns exhibited by many words. *Baker*, *writer*, and *talker* seem to pattern alike insofar as they consist of a root morpheme plus the agentive suffix *–er*. *Grocer* is superficially similar, but *groc–* is not a word and its contribution to the meaning of *grocer* is unclear. Words such as *grocer* create a dilemma for the standard approach. *Grocer* could be treated as morphologically complex because of its similarity to *baker* and *talker* and because of the analogy between *grocer–grocery* and *baker–bakery*. If *grocer* is morphologically complex, however, morphemes cannot be minimal units of meaning because *groc–* has none. (It derives from the Old French word *grossier*, "wholesale dealer," but this etymological fact is buried in the history of the language and not relevant to performance.) If the criterion that morphemes are minimal meaning-bearing units is abandoned, it leaves the basis for identifying morphemes—and therefore which words are simple and which are decomposed—unclear. Treating *grocer* as morphologically simple is equally problematic because it wrongly implies that it bears no relationship to *baker* or *writer*. Moreover, it holds that different mechanisms are involved in processing *grocer* and *baker* despite their considerable similarity.

The same issue arises in many other cases. Blueberries are blue and blackberries are black; both words fit a pattern in which a head (*berries*) is preceded by a color modifer (*blue*, *black*). Again there are deviant cases such as *cranberry* and *strawberry*, both of which resemble the other examples insofar as they refer to types of berries and superficially conform to the modifier–head structure characteristic of compounds. *Cran–* does not have an independent meaning and does not participate in other words (except for the advertising-speak neologism *cranapple*). *Strawberry*, in contrast, contains the modifier *straw*, which is a word but one whose meanings seem unrelated to strawberries. Other partial regularities are illustrated by words such as *permit*, *submit*, *remit*, and *commit*, which seem to be related insofar as they consist of a prefix and a root, but *–mit* only contributes weakly to the meanings of the words in which it occurs. Aronoff (1976) and Henderson (1985) discussed many other examples. Such cases seem to demand a less restrictive notion of morpheme than that entailed by the minimal unit of meaning idea.

In summary, derivational morphology exhibits several important characteristics. First, it is systematic: There are regularities that hold across related words such as the agentive, *–er*, cases discussed above. Second, it is productive: Knowledge of the structure of words is represented in a way that supports generalization, the comprehension, and the production of novel forms such as *geneticize*. Third, it is constrained: Some structures are clearly disallowed; thus, *frienderly* could not be a word in English. Finally, it is quasi-regular (Seidenberg & Gonnerman, 2000; Seidenberg & McClelland, 1989): There are regularities in how words are structured, but many words deviate from these central tendencies in varying degrees. Focus on clear cases that illustrate the first three characteristics has led to the development of theories in which words are represented in terms of discrete components that are accessed or activated in recognition. Such theories run up against the problem of what to do with the deviant cases. All hybrid theories assume that some words will be recognized holistically and some with component morphemes; however, they do not adequately address the question of where to draw the line. Most of the empirical studies in this area have focused on demonstrating that morphological structure is psychologically real and plays a role in processing, typically focusing on clear cases; however, there is a further question as to how the entire range of cases can be accommodated. Traditional accounts also leave many other questions unresolved, such as how the child could acquire this knowledge given the inconsistencies that exist across words. It is unclear, for example, how the child could learn that *mess* is a morpheme in *messy* but not in *message*.

It should be noted that although questions about the treatment of partial regularities arose some time ago within linguistic research on derivational morphology (e.g., Aronoff, 1976; Chomsky, 1970), no general solution to the problem has been proposed. Linguists have been committed to the idea that some words are generated by applying word formation rules, whereas others are "listed" in the mental lexicon, a view incorporated by decomposition models; however, there is little agreement about where the deviant cases fit in this dichotomy.

The Role of Semantic and Formal Factors

Much of the research on morphological processing has attempted to isolate effects of morphological structure, controlling semantic, phonological, or orthographic factors. The general consensus from these studies is that there are effects of morphological structure beyond those attributed to these other factors. However, it turns out to be quite challenging to isolate strictly morphological effects, given the correlations between morphological structure and the formal and semantic properties of words. It is very difficult to create two conditions in which the stimuli are simultaneously equated in terms of formal and semantic properties but that differ morphologically. As a consequence, researchers have tended to use a "divide and conquer" strategy in which an effect of morphological structure in one condition is compared with conditions that control either semantic or formal factors but not both. This strategy provides rather indirect evidence for effects of morphological structure, and the resulting data afford other interpretations.

To illustrate the problem, consider a classic study by Murrell and Morton (1974) that assessed priming effects for pairs of words that are morphologically related (e.g., *cars–car*) or merely similar in form (e.g., *card–car*). Morphologically related pairs produced significant priming compared with an unrelated control condition, whereas there was no reliable priming for pairs that were only formally related. Hence they concluded that the effect for *cars–car* was due to morphological relatedness. However, the effect could also be due to the fact that *cars* and *car* are semantically related, but *card* and *car* are not (see Jarvella & Snodgrass, 1974, and Stanners, Neiser, Hernon, & Hall, 1979, for similar studies).

In a subsequent study, Kempley and Morton (1982) found that regularly inflected words (e.g., *reflected–reflecting*) produced significant facilitation, whereas irregularly inflected forms (e.g., *held–holding*) did not. They concluded that "the morphemic basis of word recognition must be defined in terms of the structural, and not the semantic, properties of words" (Kempley & Morton, 1982, p. 450). Here we find the opposite confound: Although both types of word pairs are semantically related, there is less phonological overlap between the irregularly inflected forms than the regulars (e.g., *held–holding* differ in vowel quality as well as in the *–ing* suffix, whereas *reflected* and *reflecting* differ only in their suffixes). Having controlled semantic relatedness in this study, Kempley and Morton then examined priming effects for words that are only phonologically related. Phonological similarity did not produce significant priming; thus, *part* did not prime *party* nor did *deflecting* prime *reflecting*. Putting the two experiments together yielded the conclusion that priming for morphologically related pairs is not the result of either semantic or phonological relatedness. Studies by both Napps and colleagues (Fowler, Napps, & Feldman, 1985; Napps, 1989; Napps & Fowler, 1987) and by Marslen-Wilson et al. (1994) used the same strategy and reached similar conclusions. Napps (1989), for example, concluded that "morphemic priming is not the result of the convergence of semantic, orthographic, and phonological relationships but rather that morphemic relationships are represented explicitly in the lexicon" (p. 729).

There may be a problem with this general strategy, however, because the comparisons between conditions are valid only if the effects of semantic and formal factors are independent. The joint effects of these factors may be greater than the effects of either factor in isolation. Indeed, this view has found support in the area of language acquisition in which many studies have suggested that the conjunction of multiple cues can account for patterns of development that cannot be explained by single factors or even by the sum of several factors (Andersen, 1992; Bates & MacWhinney, 1982, 1987, 1989; Christiansen, Allen, & Seidenberg, 1998). We would likewise expect the relevant factors for morphology to combine in a nonlinear rather than in a merely additive way. For example, formal (e.g., phonological) similarity may produce detectable priming effects only for words that are also semantically related, as occurs in morphologically related words. The studies presented in this article address this possibility.

Morphology as an Interlevel Representation

Our approach to these phenomena is inspired by connectionist models of lexical processing and related work on inflectional morphology (e.g., Joanisse & Seidenberg, 1999; MacWhinney &

Leinbach, 1991; Plunkett & Marchman, 1991, 1993; Rumelhart & McClelland, 1986). The main reason for pursuing this approach is that it seems well suited to capturing the quasi-regular character of derivational morphology. The basic ideas are simple. Comprehending and producing language involve mappings between form (sound, spelling) and meaning. These tasks can be instantiated by connectionist networks in which there are pools of units representing spelling, sound, and meaning, and connections between them. Learning involves finding a set of weights that yields efficient, accurate translation from one code to another. In most such models, there are interlevel hidden units that allow complex relationships between codes to be represented. Morphology, on this view, reflects the systematic relationships among these codes that exist across words. *Bake*, for example, acts as a classic morpheme because it makes similar phonological, orthographic, and semantic contributions to a pool of related words. These correlations among codes are what can be encoded by a model such as the one in Figure 1. Crucially, a network trained to perform tasks such as comprehension (Phonology → Semantics) and production (Semantics → Phonology) will pick up on the correlations between codes to whatever extent they are present in the ensemble of training experiences. Thus such networks seem well suited to capture the partial regularities characteristic of many complex words.

This account is based on the same principles that have been used in models of the acquisition and use of information concerning the mapping between spelling and sound (Harm & Seidenberg, 1999; Plaut et al., 1996; Seidenberg & McClelland, 1989). There are significant parallels between the spelling–sound problem and morphology (Seidenberg & Gonnerman, 2000). The classical approach to the spelling–sound mapping is a dual-route model in which there are pronunciation rules and a separate lexical mechanism for the exceptions (Coltheart et al., 1977). The exceptions, however, exhibit partial regularities: A word such as *have*, which is standardly treated as an exception, overlaps with rule-governed forms such as *had*, *has*, and *hive*. These partial regularities, which also affect performance, are captured by connectionist models in which the weights encode different degrees of consistency in the mapping between spelling and sound. Thus, they are similar to the partial regularities in the mapping between form and meaning characteristic of derivational morphology.

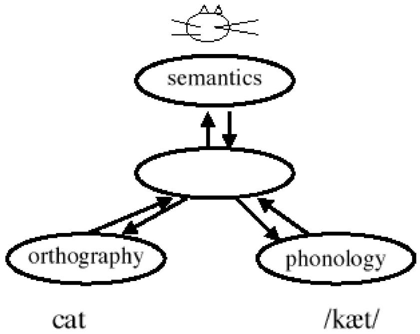

Figure 1. A distributed connectionist framework for lexical processing. The ovals represent banks of simple, neuron-like processing units. The arrows represent connections between the groups of units. Representations are patterns of activity distributed across these units, and knowledge is stored in the weights on connections between them.

Note that words in English also exhibit other correlations between form and meaning. Kelly (1992) has observed that the phonological forms of words carry probabilistic information about both meaning and grammatical function. For example, there are systematic differences between the phonological forms of nouns and verbs (Kelly, 1992) and between male and female names (Cassidy, Kelly, & Sharoni, 1999). Thus there are partial regularities in the mapping between phonology and meaning even for words, such as names, that are classically treated as morphologically simple. In our approach, these regularities are picked up by the same mechanisms that apply to "morphologically complex" words. Similarly, Mirkovic, MacDonald, and Seidenberg (2005) described a connectionist model that inferred the grammatical gender of words in Serbian from their semantic and phonological properties.

This account is a significant departure from most theories of morphological structure and processing, but it is compatible with a considerable body of research. Bybee (1985, 1995) provided extensive analyses concerning the graded character of morphological structure in several languages (including Basque, Pawnee, Yupik, Korean, and Malayalam); although she did not cast this research in terms of quasi-regularity or connectionist principles, her account is consistent with the analysis presented here and suggests that the same phenomena occur in languages other than English. Seidenberg (1987) provided evidence that statistical rather than structurally defined properties of words determine perceptual units in reading. The study examined units defined by orthographic structure, whereas the present work considers relations between phonology and semantics. Rueckl, Mikolinski, Raveh, Miner, and Mars (1997) and Li and MacWhinney (1996) also used connectionist concepts to explain effects of morphological structure on processing. Perhaps the most relevant work to that described here is the article by Rueckl et al. (1997), who used a connectionist approach to account for effects from a processing study of derivational morphology. The researchers examined long-term morphological priming in three experiments, with both masked and standard fragment completion tasks. They found that morphologically related words produce significant priming, but they argued that orthographic and phonological similarity cannot account for the results. Although these formal factors alone could not explain the experimental results, the authors pointed out that the priming effect varies in magnitude as a function of orthographic similarity. Although the authors interpreted these graded effects of orthographic similarity within a connectionist framework similar to ours, their results cannot actually distinguish between the single- and the dual-mechanism approaches because there remained morphological effects unattributable to the variation in orthographic similarity alone. Again, as in the other studies cited, the stimuli were not simultaneously controlled for semantic properties while varying surface properties.

Plaut and Gonnerman (2000) extended the approach to address issues seen as problematic for the connectionist approach. Their simulations comparing morphologically rich (e.g., Hebrew) and impoverished (e.g., English) languages demonstrated that the same principles can be successfully applied to explain morphological phenomena in typologically diverse languages. In summary, this approach suggests that morphological structure is a graded, interlevel representation that reflects the systematic though probabilistic relationships among phonological, orthographic, and semantic codes. These codes typically converge, giving rise to morphological subunits. The units are not the discrete morphemes proposed in previous theories; they encode regularities that vary in type and degree, as in the examples discussed previously.

BEHAVIORAL EXPERIMENTS

Approaching morphology from a distributed connectionist perspective leads to several specific predictions that were tested in the experiments presented below: (a) Effects attributed to morphological structure in previous studies should be predictable from semantic and phonological factors; (b) there should be intermediate cases—differences in both semantic and phonological similarity between derived words and stems form a continuum, therefore priming results between related words should be graded; and (c) effects of semantic and phonological overlap should apply to all types of morphological relationships—hence suffixed–stem, prefixed–stem, and suffixed–suffixed pairs should all prime if they are closely related in meaning and sound.

In all experiments, we used the cross-modal lexical decision task that has been used in a wide range of word processing studies, including ones on morphology (e.g., Marslen-Wilson et al., 1994). This task was chosen for three reasons. First, compared with tasks such as word naming, lexical decision promotes the processing of words to a semantic level. Second, the cross-modal aspect of the task tends to obviate repetition priming effects due to sensory overlap between prime and target (Morton, 1979). Finally, we used the task in order to be able to compare our results with those obtained in studies such as Marslen-Wilson et al., which have provided some of the strongest evidence for morpheme-based processing. In each of the four experiments, visual targets were presented immediately after the offset of the auditory prime, to maximize the likelihood of detecting effects due to different degrees of semantic similarity between the pairs. We avoided longer prime–target intervals because semantic priming tends to dissipate with longer intervals, with only the strongest effects (i.e., those for the most highly related words) being detectable (Feldman, 2000).

In Experiment 1, we examined the role of semantic similarity in processing suffixed primes and related stems, controlling for phonological similarity. We predicted larger priming effects for more highly related prime–target pairs. In Experiment 2, we examined the role of phonological overlap on processing suffixed words and their stems, controlling for semantic similarity and using only highly semantically related prime–target pairs. In Experiment 3, we tested whether the type of derivational relationship between primes and targets was important for priming to occur between semantically and phonologically related pairs by using suffixed forms for primes as well as targets. In Experiment 4, we investigated whether the same principles apply to both prefixed words and stems by examining the effect of varying degrees of semantic similarity on phonologically transparent prefixed primes and related targets.

Experiment 1: Degrees of Semantic Similarity

Experiment 1 was designed to examine priming effects for a suffixed word and its apparent root. All of the stimulus pairs exhibited the same degree of formal similarity but differed in terms of semantic similarity. Thus there were highly related pairs such as *boldly–bold*, moderately related pairs such as *lately–late*, and unrelated ones such as *hardly–hard*. The prediction was that the magnitude of priming effects would vary with the degree of semantic overlap.

Method

Participants

A group of 138 University of Southern California (USC) undergraduates completed a semantic similarity pretest. A separate group of 58 students from the same population participated in the experiment, receiving either course credit or a $5 payment. All participants learned English as their first language and used it as their primary language.

Materials

Semantic similarity pretest. A pretest was used to estimate the degree of overlap in meaning between 135 pairs of stems and words ending in a suffix (e.g., *bake*, *baker*). For each pair, the relationship between prime and target was always phonologically transparent in that the derived words contained the entire stem with no consonant or vowel changes. In addition, each longer word included a phonological segment that functions in many words as a suffix (e.g., *–er*, *–able*, *–ment*). The word pairs were alphabetized according to the stems and divided evenly into two lists. The instructions included examples of highly, moderately, and unrelated pairs and reminded participants that some words sound alike but nevertheless have quite different meanings (e.g., *ponder–pond*). Participants were then asked to rate the semantic similarity of each word pair on a scale ranging from 1 (*unrelated*) to 7 (*highly related*); they were encouraged to use the entire scale.

Mean similarity ratings were calculated for each pair of words. The data indicate that the ratings were sensitive to degrees of semantic similarity between related words. Examples are presented in Table 1 below. Responses were distributed across the entire scale: Pairs such as *message–mess* were judged unrelated ($M = 1.1$ out of 7), and pairs such as *darken–dark* were judged highly related ($M = 6.2$), but there were also many intermediate cases (e.g., *lately–late*: $M = 3.4$). The ratings were fairly evenly distributed along the scale, with no obvious discontinuities and strong

Table 1
Mean Ratings for Sample Items From Semantic Similarity Pretest

Sample word pair	Mean similarity rating
department–depart	1.60
hardly–hard	2.10
lately–late	3.40
shipment–ship	4.00
payment–pay	6.00
boldly–bold	6.60

Note. 1 = *unrelated,* 7 = *highly related.*

cross-participant agreement indicated by low standard deviations for individual pairs.[1]

Stimulus selection. The semantic similarity ratings were then used to select 84 prime–target pairs, falling into three conditions with 28 items in each (Table 2). Items in the low semantic set (Condition 2) had mean ratings of <3; these were items such as *hardly–hard*. For the moderate semantic set (Condition 3), the ratings were ≥ 3 and <5 (e.g., *lately–late*), and for the high semantic set (Condition 4) the ratings were ≥ 5 (e.g., *boldly–bold*). Different suffixes were represented approximately equally across conditions so that responses were not biased in a particular condition by the type of suffix (i.e., the distributions of words ending in *–er*, *–ly*, *–age*, etc., were very similar across conditions).

Two additional conditions were included to examine pairs that exhibit phonological (form only) or semantic (semantic only) similarity in the absence of real or pseudomorphological overlap. The form only set (Condition 1) consisted of 28 prime–target pairs that were phonologically transparent but semantically unrelated (e.g., *spinach–spin*); these pairs were similar to those in the low semantic condition (e.g., *hardly–hard*), except that the form only test primes do not end in phonological segments that function as suffixes; for example, the *–ach* in *spinach* occurs as rime of only a few other English words (e.g., *stomach*, *detach*), does not carry a systematic meaning, and varies in pronunciation. Another example is the *–ow* in *shallow*, which appears as a syllable at the end of several words but does not recur with a consistent meaning (*yellow*, *wallow, fellow, billow*). In contrast, although the items in the low semantic condition would not be considered suffixed on most linguistic accounts, the endings do function as suffixes in other words (e.g., the *–er* in *corner* vs. *baker* or *–ment* in *pigment* vs. *government*). In addition, *corn* occurs in other words such as *cornish*, *cornice, cornea,* and *corning*, but these forms are unrelated to the meaning of *corner*; thus, *corner* is not part of a morphological paradigm with *corn* as its stem. Such words have been termed *pseudoaffixed* (Taft, 1979).

In our account, morphological structure is graded: It reflects the extent to which words exhibit similar sound–meaning mappings, regardless of historical or morphological relationships. It is important to include both form only and low semantic word types because the use of primes with embedded words but with no pseudosuffixes could provide evidence for an independent effect of morphology. For example, finding significant facilitation effects for the low semantic (e.g., *corner–corn*) condition but not for form only (e.g., *spinach–spin*) would indicate that a factor other than semantic or phonological similarity (perhaps morphological structure) contributed to the priming effects, because these conditions are matched on semantic and phonological similarity measures. The absence of differential effects between these two conditions would be consistent with the hypothesis that semantic and phonological similarity, but not morphological structure, underlie any facilitation of targets following related primes. Because sufficient semantic similarity is hypothesized to be crucial for any priming to occur, neither of these conditions should produce facilitation effects because both conditions are low in mean semantic similarity (see Table 2).

Finally, in the semantic only set (Condition 5), the 28 word pairs were synonyms (thus highly semantically related) with no phonological similarity (e.g., *idea–notion*). Sample stimuli for each condition and the mean similarity ratings are shown in Table 2.

For each of the 140 test primes (five conditions of 28 items each), a control prime was selected to match the test prime in frequency, number of syllables, and part of speech (see Table 3 for mean [Kučera & Francis, 1967] frequency values). Test and control primes were neither phonologically nor semantically related. In addition, to avoid experiment-specific response strategies, we included 140 nonword fillers, some phonologically related (e.g., *slither–slith*) and others not (e.g., *basil–grook*). The phonologically related nonwords consisted of two types, ones with pseudoaffixes (e.g., *slither–slith*) and ones without recognizable affixes (e.g., *bishop–bish*). The *slither–slith* pairs resemble *corner–corn* pairs, in that both end in a syllable, *–er*, that functions as a suffix in many words (e.g., *teacher*, *baker*) but does not happen to perform that function in this case. The *bishop–bish* pairs resemble *spinach–spin* pairs in that the endings *–op* or *–ach* never function as suffixes in any English words. The proportion of phonologically related word types to phonologically unrelated words was matched for the nonwords to the proportions of overlapping words in the real word stimuli. This was done to ensure that participants could not develop a strategy whereby all phonologically related words or all words that ended in suffixes could be correctly responded to as real words. The items were divided into two lists, one with the test–target pair (e.g., *cowardly–coward*) and the other with the control–target pair (e.g., *demented–coward*); each participant saw the stimuli from only one list. Two separate, pseudorandom orders of all the items were generated to create a total of four lists. All of the test and control primes were digitally recorded by a female, native English speaker.

Table 2
Sample Stimuli and Mean Similarity Ratings From Experiment 1 (Graded Semantic Similarity—Suffixes)

Condition	Prime–Target example	Mean semantic similarity *M*	*SD*
Form only	*spinach–spin*	1.20	1.06
Low semantic	*hardly–hard*	1.90	1.33
Moderate semantic	*lately–late*	3.90	1.87
High semantic	*boldly–bold*	6.10	1.19
Semantic only	*idea–notion*	6.00	1.52

Note. 1 = *unrelated,* 7 = *highly related.*

Procedure

Participants were tested individually in a quiet room, seated in front of a Macintosh computer with a 14-in. color monitor. Participants listened to primes played on a high-quality speaker placed next to the testing computer. Lexical decisions were indicated by pressing a button on a button box. Rapid and accurate responses were encouraged. PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) software was used to present stimuli and to record responses.

---

[1] For all of the experiments reported in this article, the means and standard deviations for the items in the semantic similarity pretests are available from Laura M. Gonnerman upon request. Reaction times are also available upon request.

Table 3

Mean Lexical Decision Latencies (in Milliseconds), Frequencies, and Error Rates for Experiment 1 (Graded Semantic Similarity—Suffixes)

		Reaction time		Control frequency			Reaction time			Test frequency		
Condition	Example	*M*	*SD*	*M*	*SD*	Error rate	*M*	*SD*	Error rate	*M*	*SD*	Priming effect
Form only	*spinach–spin*	649	161.60	11.60	3.09	.05	668	140.80	.10	14.00	4.70	−19
Low semantic	*hardly–hard*	607	125.30	22.10	5.50	.02	631	127.70	.02	25.20	9.60	−24
Moderate semantic	*lately–late*	588	102.60	26.30	7.03	.03	569	106.90	.02	31.50	9.10	19^{*}
High semantic	*boldly–bold*	613	107.70	41.10	17.50	.02	573	107.70	.02	49.30	18.40	40^{*}
Semantic only	*idea–notion*	593	93.80	25.70	8.70	.01	580	101.40	.01	34.60	9.50	13^{*}

* $p < .05$.

At the start of each trial, a fixation point (three asterisks) was displayed at the center of the screen for 1 s, followed by presentation of the auditory prime. Immediately after the offset of the prime word, the target was displayed on the screen for 200 ms. After the participant responded, there was a 500-ms delay before the next trial began. All targets were presented as lowercase letters in 14-point type with a sans serif font. The trial ended when a button was pushed indicating the response. To ensure that participants were attending to the recorded primes, the instruction "Please repeat the word you just heard" was presented on the computer screen after 15% of the trials. Participants' responses were recorded by the experimenter. Participants then pressed either of the response buttons to proceed to the next trial.

Participants were given 20 practice items, followed by 4 warm-up items before presentation of the 280 word and nonword test stimuli. Fifteen participants were tested on List 1, 15 on List 2, 14 on List 3, and 14 on List 4. The experiment took approximately 25 min to complete, including practice, warm-up, and test trials.

Results and Discussion

Data from 1 participant were excluded because his error rate was excessive (over 30% errors). In addition, 9 items were excluded because of high error rates (over 50%). Three of these items were from Condition 1, 3 from Condition 2, and 3 from Condition 3. This left a total of 57 participants and 131 items.

Trials on which participants made an error (3.0%) were excluded from the latency analyses as were outliers (responses greater than 2,000 ms or less than 200 ms: 0.8%). The decision latencies were entered into an analysis of variance (ANOVA) with the variables prime type (test or control) and condition (the five types of prime–target relations: form only, low semantic, moderate semantic, high semantic, and semantic only). All means presented are based on analyses by participants.[2] Summary data are presented in Table 3. Mean decision latencies and error rates for the nonword items are presented in the table in Appendix A.

The main effect of prime type was not significant, showing facilitation in some conditions and interference in others: $F(1, 56) < 1$. The main effect of condition was significant, $F(4, 224) = 27.9$, $MSE = 4285.00$, $pr\eta^2 = .33$, $p < .05$.[3] Although the stimuli were matched across conditions, decision latencies were longer in the form only condition in both the test and control conditions, indicating that some of these items may have been less familiar to USC undergraduates. Finally, the Prime Type × Condition interaction was significant, $F(4, 224) = 7.9$, $MSE = 2623.00$, $pr\eta^2 = .12$. The differences between test and control primes by condition are shown in Table 3.

There were numerically negative effects in the form only and low semantic similarity conditions, −19 and −24 ms, respectively. Planned comparisons showed that neither difference was significant ($F < 1$, for both) because the priming effects in these conditions were not consistent across participants (i.e., the Participant × Effect interactions were large). These data suggest that morphological structure does not produce reliable priming effects in the absence of semantic overlap. In the other experiments in this article, we report replications of these conditions that suggest that the absence of the effect is not a Type II error. In addition, retrospective power analyses indicate that this experiment had a sensitivity of .70 or more for detecting priming effects in the form only and low semantic conditions on the basis of the effect sizes and standard deviations observed (Buchner, Faul, & Erdfelder, 1992; Erdfelder, Faul, & Buchner, 1996).

The moderate and highly semantically related conditions yielded significant priming effects, with the magnitude of the effect determined by the degree of similarity. Moderately related words (e.g., *lately–late*) primed about half as much (19 ms vs. 40 ms) as highly related words (e.g., *boldly–bold*). The effects in both of these conditions were significant: moderate, $F(1, 56) = 4.51$, $MSE = 2266.00$, $pr\eta^2 = .07$; high, $F(1, 56) = 19.64$, $MSE = 2366.00$, $pr\eta^2 = .26$. The semantic only condition (pairs such as *idea–notion*) also yielded a significant priming effect of 13 ms, $F(1, 56) = 5.72$, $MSE = 964.00$, $pr\eta^2 = .09$. The stimuli in this condition were as semantically related as in the high semantic condition but the priming effect was smaller, presumably due to the absence of phonological overlap.

The results from this experiment indicate that, with the degree of phonological overlap equated, the magnitude of the priming effect

---

[2] Item analyses were also calculated. They are not reported here because the items served as their own controls in the experiment, the effects are replicated across different sets of items, and the item analyses are consistent with the participant analyses, so they are not reported here. *F*2 values are available from Laura M. Gonnerman upon request.

[3] For the statistical tests reported in this article, we use the following conventions: $\alpha = .05$, and p values are reported only when $F > 1$ or $t > 1$ and they are nonsignificant, in which case exact p values are reported.

is related to the degree of semantic similarity. To examine this further, we conducted a regression analysis by using rated semantic similarity to predict the differences between test and control conditions for all items in the low, moderate, and high semantic conditions. The data from the semantic only condition were excluded because they were not comparable to the other conditions in terms of formal overlap. In addition, data from the form only condition were excluded because these items do not end with a putative suffix. Semantic similarity in the form only condition is consistently low, whereas there is a gradation of similarity, albeit on the low end of the scale, for the low semantic condition. Semantic similarity was a significant predictor of priming effects ($r = .41$, $p < .001$), indicating that participants are sensitive to subtle differences in the similarity of meanings of pairs of words and that the ratings provide reliable information about degree of semantic similarity as reflected in the priming effects. To further examine the graded facilitation based on semantic similarity, we also conducted a comparison of the moderate and high semantic conditions by using an ANOVA comparing Condition (moderate vs. high semantic) × Prime Type (test or control). This analysis approached significance, suggesting that high semantic items produced greater facilitation than did moderate semantic items, $F(1, 56) = 3.79$, $MSE = 1996.00$, $pr\eta^2 = .06$. Taken together, these analyses suggest graded effects based on semantic similarity for morphologically related word pairs.

Data concerning trials on which participants made errors were also analyzed. Items that were trimmed as outliers were not included in the error analysis. This procedure was followed for all experiments. The error rates were entered into an ANOVA with the variables of prime type (test or control) and condition (the five types of prime–target relations). Error rates by condition are shown in Table 3.

The main effect of prime type approached significance, $F(1, 56) = 3.93$, $MSE = 0.0024$, $pr\eta^2 = .07$, indicating that it was generally more difficult to respond correctly to the targets when they were preceded by test primes compared with unrelated control primes. There was a significant main effect of condition, $F(4, 224) = 34.6$, $MSE = 0.0022$, $pr\eta^2 = .38$, because one condition (form only) produced more errors in both the test and control conditions than in the others. There was also a significant Prime Type × Condition interaction, $F(4, 224) = 10.9$, $MSE = 0.0018$, $pr\eta^2 = .16$. This interaction indicates that there was a higher error rate in the primed compared with the control condition but only for the form only items. Thus the priming effects seen in the reaction time analyses in the other conditions are not compromised by a speed–accuracy trade-off.

The overall pattern of results for this experiment indicates that priming effects between pairs of related words can be predicted by the degree of semantic similarity between the prime and the target. Furthermore, the priming effects for suffixed primes (e.g., *baker*) and related targets (e.g., *bake*) are graded, such that priming increases with increases in semantic similarity.

Experiment 2: Degrees of Phonological Similarity

The results from Experiment 1 suggest that the degree of semantic relatedness modulates the priming effects observed with formally related words. In this experiment, we investigated the role of sound overlap between complex words and stems. Marslen-Wilson et al. (1994) conducted a study in which pairs that are phonologically related but semantically and morphologically unrelated (e.g., *tinsel–tin*) did not produce reliable priming effects, whereas reliable priming was observed for morphologically related pairs such as *friendly–friend*. They therefore concluded that the priming for morphologically related pairs was not due to phonological overlap. We were interested in the possibility that the effects of phonological similarity depend on the degree of semantic relatedness. If morphological structures arise from the convergence of semantics and phonology, we might expect these factors to show interactive effects. In particular, Experiment 2 tested the hypothesis that amount of phonological overlap modulates priming effects for words that are highly semantically related.

Method

Participants

A group of 93 USC undergraduates participated in a semantic similarity pretest. A separate group of 51 participants from the same population participated in the experiment, receiving either course credit or a $5 payment. All participants learned English as their first language and used it as their primary language. None of the participants in either the pretest or the experiment itself had participated in Experiment 1.

Materials

Phonological similarity metric. It seems clear that complex words differ in the extent to which they phonologically overlap with their stems; for example, *triumph* and *triumphant* have a more transparent phonological relationship than do *sign* and *signal*. To assess how complex words vary along this dimension, we created a phonological similarity metric. This metric involved the following two assumptions: (a) A vowel change between a stem and a derived form creates more difference in sound than does a consonant change; and (b) distance is a function of the number of differing phonemes—thus, for example, a consonant change accompanied by a vowel change creates more distance than either a consonant change or a vowel change alone. These two assumptions generated four similarity conditions, ordered from most to least phonologically similar: (a) no change, where the derived form contains the complete stem without phonological modification (e.g., *acceptable–accept*); (b) consonant change, where there is a change in a consonant of the stem in the derived form (e.g., *absorption–absorb*); (c) vowel change, where the derived form differs from the stem in vowel quality only (e.g., *criminal–crime*); and (d) consonant-plus-vowel change, where the stem in the suffixed form differs in both a consonant and a vowel from the stem in its simple form (e.g., *introduction–introduce*).

Semantic similarity pretest. To ensure that any effects obtained in the lexical decision experiment were due to differences in the degree of phonological similarity between primes and targets, it was necessary to control for semantic similarity between primes and targets across conditions. To obtain a set of prime and target pairs that were matched in degree of semantic similarity, candidate word pairs of different levels of phonological overlap were included in a semantic similarity pretest.

For each of the first three phonological similarity categories described above, 56 word pairs were chosen. Items in the last category, consonant-plus-vowel change, were most difficult to find; therefore, only 46 items of that type were included. Four lists were created, one for each phonological similarity category. In addition to the phonologically related word pairs, 60 filler items were also included, which fell into four classes of 15 pairs each: (a) morphologically related but semantically distant word pairs (e.g., *succession–successful*); (b) both morphologically and semantically unrelated pairs (e.g., *violence–violin*); (c) both morphologically and semantically related pairs (e.g., *boyish–boyhood*); and, finally, (d) synonyms (e.g., *porpoise–dolphin*). These 60 filler items were included on each list to promote the use of the full range of the similarity scale. The procedure was the same as for Experiment 1, except that participants were asked to use a scale ranging from 1 (*unrelated*) to 9 (*highly related*). Mean similarity ratings were calculated for each pair of words.

Stimulus selection. The semantic similarity ratings were used to select 168 prime–target pairs, falling into six conditions with 28 items each (Table 4). Conditions 1–4 are based on the four levels of the phonological similarity metric described earlier: no change, consonant change, vowel change, and consonant-plus-vowel change.[4] It is important that test primes and targets in these four conditions were highly semantically similar, and the items are equated for semantic similarity across conditions. As in Experiment 1, semantically unrelated word pairs (e.g., *accordion–accord*) were included in a low semantic set (Condition 5), and phonologically unrelated synonyms (e.g., *porpoise–dolphin*) were included in a semantic only set (Condition 6).

For each of the 168 test primes, a control prime was selected to match the test prime in frequency, number of syllables, and part of speech. Control primes were neither phonologically nor semantically related to the targets. In addition, to avoid experiment-specific response strategies, 168 varied nonword fillers were included. There were 70 phonologically related primes and nonword targets that paralleled the differences in phonological similarity of Conditions 1–4, including pairs with no change (e.g., *territory–territ*), ones with a consonant change only (e.g., *foundation–foundate*), ones with a vowel change only (e.g., *marital–marite*), and ones with both a consonant and a vowel change (e.g., *struggle–struge*). There were 98 phonologically unrelated nonword targets (e.g., *boomerang–jaulic*), with primes that were matched in grammatical category, frequency, and number of syllables to the word primes to minimize any strategies that participants might develop on the basis of those stimuli characteristics (see Table 5 for mean [Kučera & Francis, 1967] frequency values). The items were divided initially into two counterbalanced lists, one with the test–target pair (e.g., *division–divide*) and the other with the control–target pair (e.g., *manager–divide*); thus, each participant saw each target only once, preceded either by the corresponding test or control prime. Half of the test–target pairs and half of the controls appeared on each list. Two pseudorandom orders of all the items were generated to create a total of four lists. All of the test and control primes were digitally recorded by a female, native English speaker.

Table 4
Sample Stimuli and Mean Similarity Ratings From Experiment 2 (Graded Phonological Similarity)

Condition	Prime–target example	*N*	Semantic similarity *M*	*SD*
No change	*acceptable–accept*	28	7.40	1.71
Consonant change	*absorption–absorb*	28	7.60	1.63
Vowel change	*criminal–crime*	28	7.50	1.54
Consonant-plus-vowel change	*introduction–introduce*	28	7.40	1.61
Low semantic	*accordion–accord*	28	2.00	0.81
Semantic only	*porpoise–dolphin*	28	7.60	0.85

Note. 1 = *unrelated,* 9 = *highly related.*

Procedure

The procedure was the same as described in Experiment 1. Participants were given 20 practice items, followed by 5 warm-up items before presentation of the 336 word and nonword test stimuli. Thirteen participants were tested on Lists 1, 2, and 3, and 12 were tested on List 4. The experiment took approximately 30 min to complete, including practice, warm-up, and test trials.

Results and Discussion

Trials on which participants made an error (2.5%) were excluded from the latency analysis as were outliers (responses greater than 2,000 ms or less than 200 ms; 0.5%). The decision latencies were entered into an ANOVA with the variables prime type (test or control) and condition (the six types of prime–target relations: no change, consonant change, vowel change, consonant-plus-vowel change, low semantic, and semantic only). All means presented are based on participant analyses. Summary data are presented in Table 5. Mean decision latencies and error rates for the nonword items are presented in the table in Appendix B.

The main effect of prime type was significant, $F(1, 50) = 78.9$, $MSE = 2660.00$, $pr\eta^2 = .61$, indicating that responses to the target items were faster overall following the test primes compared with the control primes. The main effect of condition was significant, $F(5, 250) = 34.5$, $MSE = 2411.00$, $pr\eta^2 = .41$. Although the stimuli were matched across conditions, decision latencies were longer in the low semantic condition in both test and control conditions, indicating that some of these items may have been less familiar to USC undergraduates. Finally, the Prime Type × Condition interaction was significant, $F(5, 250) = 8.4$, $MSE = 2156.00$, $pr\eta^2 = .14$. The differences between test and control primes in individual conditions are shown in Table 5.

The low semantic condition, in which primes and targets were unrelated in meaning, yielded a 14-ms negative effect that did not reach significance, $F(1, 50) = 1.6$, $MSE = 2892.00$. This result is similar to that found for the low semantic condition in Experiment 1, which also yielded a small but statistically nonsignificant negative effect. The semantic only condition (pairs such as *porpoise–dolphin*) yielded a significant priming effect of 40 ms, $F(1, 50) = 19.0$, $MSE = 2122.00$, $pr\eta^2 = .2750$.

Planned comparisons for Conditions 1–4 showed significant priming effects for each degree of phonological similarity. The

---

[4] Stress-shifted items were not excluded from the stimuli and are represented across the conditions.

Table 5

Mean Lexical Decision Latencies (in Milliseconds), Frequencies, and Error Rates for Experiment 2 (Graded Phonological Similarity)

Condition	Example	Reaction time		Control frequency		Error rate	Reaction time		Test frequency		Error rate	Priming effect
		M	*SD*	*M*	*SD*		*M*	*SD*	*M*	*SD*		
No change	*acceptable–accept*	623	109.70	21.40	5.80	.02	576	106.60	5.70	.01	26.40	47*
Consonant change	*absorption–absorb*	662	112.60	32	11.70	.02	597	107.30	5.90	.01	24	65*
Vowel change	*criminal–crime*	656	114.90	33.90	18.80	.02	608	122.70	14.20	.01	31.50	48*
Consonant-plus-vowel change	*introduction–introduce*	674	111.10	27	7.90	.04	639	123.70	6.60	.02	21	35*
Low semantic	*accordion–accord*	677	121.40	36	9.60	.06	691	116.90	10.50	.07	36	−14
Semantic only	*porpoise–dolphin*	661	112.40	16	5.20	.03	621	115.80	9.20	.01	32	40*

* $p < .05$.

differences between test and control means were significant in all four cases: no change, $F(1, 50) = 39.0$, $MSE = 1477.00$, $pr\eta^2 = .44$; consonant change, $F(1, 50) = 38.5$, $MSE = 2760.00$, $pr\eta^2 = .44$; vowel change, $F(1, 50) = 33.0$, $MSE = 1820.00$, $pr\eta^2 = .40$; consonant-plus-vowel change, $F(1, 50) = 13.5$, $MSE = 2368.00$, $pr\eta^2 = .21$.

We predicted the greatest priming effect for the no change condition (most phonologically related), the smallest effect for the consonant-plus-vowel-change condition (least phonologically related), and intermediate effects for the consonant change and vowel change conditions (both moderately related). This prediction was upheld for the consonant, vowel, and consonant-plus-vowel-change conditions. However, the no change condition primed less than the consonant change condition (47 vs. 65 ms), although the difference between the conditions was not significant. The smaller effect in the no change condition may be because many of the derived words in that set are resyllabified compared with their stems. For example, the /b/ in *absorbent* is in the onset of the third syllable, *–bent*, but in the coda of the second syllable in *absorb*. Perhaps *absorption*, even though it entails a consonant change, should be considered more similar to *absorb* than to *absorbent* because *absorption* retains the syllable structure of *absorb*, with the stop /p/ in the final coda of the stem. In any case, the most important prediction of a general decrease in magnitude of priming effects as stimulus pairs decrease in phonological similarity was supported, especially in that consonant change items (e.g., *absorption–absorb*) prime significantly more than consonant-plus-vowel-change items (e.g., *decision–decide*), $F(1, 50) = 6.44$. In addition, a test for a linear trend including consonant change, vowel change, and consonant-plus-vowel-change (in that order in keeping with our predictions) was significant, $F(1, 49) = 10.44$. A more sensitive measure of phonological similarity might be better able to detect subtle differences in priming magnitude across decreases in phonological similarity.

Data concerning trials on which participants made errors were also analyzed. The error rates were entered into an ANOVA with the variables of prime type (test or control) and condition (the six types of prime–target relations). Error rates by condition are shown in Table 5.

There was a significant main effect of prime type, $F(1, 50) = 9.86$, indicating that it was generally more difficult to respond correctly to the targets when they were preceded by unrelated control primes compared with test primes, although the magnitude of the difference was small (3% error for targets following control primes and 2% error for targets following test primes). There was a significant main effect of condition, $F(5, 250) = 23.42$, because one condition (low semantic) produced more errors in both test and control conditions than did the others. There was also a significant Prime Type × Condition interaction, $F(5, 250) = 2.24$.

Overall, the pattern of priming in this experiment indicates that phonological similarity is also a significant predictor of priming effects. Thus, for highly semantically similar pairs of words, the more phonologically similar the prime is to the target, the greater the facilitation.

Experiment 3: Role of Morphological Type

The model proposed by Marslen-Wilson et al. (1994) relied heavily on the result of a priming experiment that showed significant facilitation for related derived and stem pairs, such as *government–govern*, but no significant priming for pairs of related suffixed words, such as *government–governor*. This result led Marslen-Wilson et al. to propose a model of the mental lexicon in which stems are linked to suffixes, and suffixes inhibit one another. Experiment 3 included a replication of the critical suffixed–suffixed condition. Our hypothesis was that such pairs will yield priming if both words are sufficiently similar in meaning and sound.

Method

Participants

Forty-seven USC undergraduates completed the semantic similarity pretest. A separate group of 51 students from the same population participated in the experiment, receiving either course credit or a $5 payment. None of the participants had participated in either Experiment 1 or Experiment 2. All participants learned English as their first language and used it as their primary language.

Materials

Semantic similarity pretest. A pretest was used to estimate the degree of overlap in meaning between 202 pairs of words: 68 pairs of suffixed words (e.g., *sainthood–saintly*), 28 synonyms (e.g., *sorcery–magic*), 28 semantically unrelated but phonologically similar pairs (e.g., *catacomb–catalog*), and 78 morphologically unrelated filler items. The word pairs were divided evenly into two lists. Participants were asked to rate the semantic similarity of each word pair by using a scale ranging from 1 (*unrelated*) to 9 (*highly related*). Participants were given examples of highly related as well as unrelated pairs along with sample ratings. The instructions were the same as for Experiment 1. Mean similarity ratings were calculated for each pair of words.

Stimulus selection. The semantic similarity ratings were used to select 60 prime–target pairs, falling into three conditions (see Table 6). Items in the high phonology/moderate semantic set (Condition 1) were phonologically transparent (e.g., *useful–useless*) but were rated as only moderately similar semantically. Items in the high phonology/high semantic set (Condition 2) were both phonologically transparent and highly semantically similar (e.g., *scientific–scientist*). Items in the low phonology/high semantic set (Condition 3) were less phonologically similar but were highly semantically related (e.g., *observation–observant*).

Many of the items in the high phonology/moderate semantic condition were opposites, such as *harmless–harmful*, or exhibited differences in thematic role, for example *drinker–drinkable*. Participants consistently gave these types of word pairs lower semantic similarity ratings. It was therefore unclear whether test primes in this condition would facilitate reaction times to their related targets. Items in the high phonology/high semantic condition were expected to prime, even though both primes and targets were suffixed forms. The low phonology/high semantic condition consisted of suffixed pairs with characteristics similar to those used by Marslen-Wilson et al. (1994). The phonological relatedness of these pairs was not transparent: There were changes in either vowel or consonant quality in the stems of the two suffixed forms (e.g., *observation–observant*). In addition, the items in this condition were also somewhat less semantically related than items in high phonology/high semantic condition. Because the low phonology/high semantic items were less related both phonologically and semantically than were pairs in the high phonology/high semantic condition, we expected them to yield smaller effects, as in the Marslen-Wilson et al. experiment, which yielded an 11-ms effect that was not statistically significant. Sample stimuli for each condition and the mean similarity ratings are shown in Table 6 below.

For each of the 60 test primes, a control prime was selected to match the test prime in frequency, number of syllables, and part of speech (see Table 7 for mean [Kučera & Francis, 1967] frequency values). Test and control primes were neither phonologically nor semantically related. In addition, 90 filler prime–target pairs were included: 15 semantically unrelated pairs, where each word ended in a suffix (e.g., *casually–casualty*); 15 semantically unrelated pairs that did not end in suffix-like phonological segments (e.g., *catacomb–catalog*); and 15 synonyms (e.g., *sorcery–magic*). Because both Experiments 1 and 2 provided evidence concerning the priming effects for these types of pairs, they were treated as fillers and thus matched control primes were not included for them. Instead, an additional 45 completely unrelated (i.e., morphologically, semantically, and phonologically unrelated) prime–target pairs were included (e.g., *admiration–exclusive*). These last primes were matched for part of speech, frequency, and number of syllables with the primes in Conditions 1–3. This yielded a total of 150 word pairs.

The nonword stimuli were 150 pairs of three types: 30 phonologically overlapping with suffixes (e.g., *respectful–respection*), 30 phonologically overlapping without suffixes (e.g., *stylist–styleem*), and 90 phonologically unrelated (e.g., *optimal–brovian*). The primes for the nonword targets were matched in grammatical category, frequency, and number of syllables to the primes for the word targets to minimize any strategies that participants might develop on the basis of these stimulus characteristics. The items were divided into two counterbalanced sets, with the test–target pair (e.g., *saintly–sainthood*) on one list and the control–target pair on the other. Half the items of each type appeared on each list. Two pseudorandom orders of all the items were generated to create four lists. Each participant was presented with one list. The test and control primes were digitally recorded by a female native English speaker.

Procedure

The procedure was the same as described in Experiment 1, except that the targets were displayed for 250 rather than 200 ms. The extra time was considered necessary to read the somewhat longer suffixed targets. Participants were given 20 practice items, followed by 4 warm-up items before presentation of the 300 word and nonword test stimuli. Thirteen participants were tested on Lists 1, 2, and 3, and 12 were tested on List 4. The experiment took approximately 30 min to complete, including practice, warm-up, and test trials.

Results and Discussion

Trials on which participants made an error (3.4%) and extreme values (responses greater than 2,000 ms or less than 200 ms; 0.3%) were excluded from the latency analyses. The decision latencies were entered into an ANOVA with the variables of prime type (test or control) and condition (the 3 types of prime–target relations: high phonology/low semantic; high phonology/high semantic; low

Table 6
Sample Stimuli and Mean Similarity Ratings From Experiment 3 (Suffixed–Suffixed Pairs)

Condition	Prime–Target example	*n*	Semantic similarity *M*	Semantic similarity *SD*
High phonology, moderate semantic	*useful–useless*	15	4.70	2.44
High phonology, high semantic	*saintly–sainthood*	30	7.70	1.49
Low phonology, high semantic	*observation–observant*	15	7.40	1.69

Table 7
Mean Lexical Decision Latencies (in Milliseconds), Frequencies, and Error Rates for Experiment 3 (Suffixed–Suffixed Pairs)

		Reaction time		Control frequency			Reaction time		Test frequency			
Condition	Example	*M*	*SD*	*M*	*SD*	Error rate	*M*	*SD*	*M*	*SD*	Error rate	Priming effect
High phonology, low semantic	*useful–useless*	624	92.70	17.40	9.60	.02	628	130.90	5.10	.01	12.20	−4
High phonology, high semantic	*saintly–sainthood*	655	95.10	12.80	5.60	.03	621	92.50	4.40	.02	12.80	34*
Low phonology, high semantic	*observation–observant*	652	123.60	17.60	5.10	.01	638	113.60	5.20	.01	24.80	14

* $p < .05$.

phonology/high semantic). All means presented are based on participant analyses. Summary data are presented in Table 7. Mean decision latencies and error rates for the nonword items are presented in the table in Appendix C.

The main effect of prime type was significant, $F(1, 50) = 4.60$, $MSE = 3423.00$, $pr\eta^2 = .08$. The main effect of condition approached significance, $F(2, 100) = 2.76$, $MSE = 3352.00$, $pr\eta^2 = .05$, $p = .07$. Finally, the Prime Type × Condition interaction was significant, $F(2, 100) = 3.07$, $MSE = 2868.00$, $pr\eta^2 = .0578$. The differences between test and control primes in individual conditions are shown in Table 7.

Planned contrasts were conducted comparing the control and test latencies for each condition. For the high phonology/low semantic pairs there was a nonsignificant effect of −4 ms ($F < 1$). Thus, hearing *useful* does not facilitate responses to *useless*. The items in the high phonology/high-semantic condition yielded a significant 34-ms facilitation effect, $F(1, 50) = 21.0$, $MSE = 1349.00$, $pr\eta^2 = .2956$, $p < .001$. Finally, for the low phonology/high semantic items there was a slight facilitation effect, 14 ms, that did not reach significance, $F(1, 50) = 1.3$, $MSE = 3601.00$, $p = .26$. It is interesting to note that the result in the low phonology/high semantic condition is very similar to that obtained by Marslen-Wilson et al. (1994), who found an 11-ms facilitation effect for their suffixed prime–target pairs that also failed to reach significance. There are two possible interpretations of these effects: (a) the semantic similarity of the items in the low phonology/high semantic condition was not high enough to reach a threshold level needed to produce significant priming; or, more likely, (b) the marginal facilitation effect reflects an intermediate level of priming, which might reach significance with a larger number of participants and items. A power analysis indicates that this experiment had a sensitivity of .51 for detecting a priming effect in this condition, based on the effect size and variance observed (Buchner et al., 1992; Erdfelder et al., 1996). Because power is based in part on effect size, this result supports our interpretation of the effects, namely, the magnitude of the effect is smaller because the items are less semantically similar and thus produce less facilitation. Obtaining a significant effect in a condition in which the effect is necessarily smaller would require more participants and/or items.

Analyses of the error rates for word targets in the various conditions were also carried out. The error rates were entered into an ANOVA with the variables of prime type (test or control) and condition (the three types of prime–target relations). The error rates were very low with no differences between the targets following control and test primes in Conditions 1–3. There were no significant effects of prime type, condition, or the Prime Type × Condition interaction ($F < 1$). The error rates for each condition are shown in Table 7.

The results indicate that hearing a suffixed word facilitates lexical decision to another suffixed word when the words are sufficiently similar in both meaning and sound. Furthermore, the results underscore the importance of considering the joint rather than the independent effects of semantic and phonological similarity. In the high phonology/low semantic and low phonology/high semantic conditions, in which the stimuli were highly related on only one dimension, there were no significant priming effects. Only in the high phonology/high semantic condition, in which primes and targets were both highly semantically and phonologically related was there significant facilitation. These results are consistent with the Marslen-Wilson et al. (1994) results insofar as the stimuli that were comparable to their word pairs did not yield significant priming. However, contrary to their conclusion, suffixed–suffixed word pairs do prime if they exhibit sufficient phonological and semantic overlap.

Experiment 4: Role of Morphological Type—Prefixed–Stem Pairs

Experiment 4 examined whether there are differences in processing for prefixed versus suffixed words. Some researchers have proposed that the two types of words involve different processing mechanisms; one mechanism is invoked when the affix is encountered before the stem, but a different mechanism is invoked when the affix follows the stem (e.g., Andrews, 1986; Colé, Beauvillain, & Segui, 1989; Marslen-Wilson et al., 1994). Prefixed and suffixed words differ systematically in ways that may affect performance on tasks such as spoken word recognition; for example, prefixed words may take longer to identify because the prefix initially activates a large cohort of candidates, whereas the cohort associated with a suffixed word is smaller. The prediction derived from our account, however, is that semantic and phonological factors should affect the processing of both types of affixed words. Experiment 4 addressed whether the degree of semantic relatedness predicts the magnitude of priming effects in prefixed–stem pairs as it did in Experiment 1 for suffixed stimuli.

Method

Participants

One hundred twenty-four Carnegie Mellon University undergraduates completed the semantic similarity pretest. A separate group of 42 students from the same population participated in the experiment for course credit. All participants learned English as their first language and used it as their primary language.

Materials

Semantic similarity pretest. A semantic similarity pretest was used to determine the degree of overlap in meaning between pairs of stems and prefixed words. The phonological relationship between the words was always transparent, in that the prefixed words contained the entire stem with no consonant or vowel changes (e.g., *preheat* contains *heat*). Instructions were similar to the pretest described in Experiment 1.

Mean similarity ratings were calculated for each pair of words. As with the suffixed words, participants used the entire scale: items such as *rehearse–hearse* were judged dissimilar ($M = 1.3$ out of 9), others were judged as highly semantically similar (e.g., *midnight–night*, $M = 7.9$), and there were intermediate cases (e.g., *premature–mature*, $M = 5.7$). The ratings were fairly evenly distributed along the scale and there was strong cross-participant agreement, as indicated by low standard deviations for individual pairs. Note that for each prefix (e.g., *mid–, re–, pre–, be–, de–*) ratings were distributed along the scale as well: For example, *en–* was represented at low (*enchant*), moderate (*enrich*), and high (*enslave*) similarity levels.

Stimulus selection. The semantic similarity ratings were then used to select 84 prime–target pairs, falling into three conditions of 28 items each (Table 8). Items in the low semantic set (Condition 2) were rated less than 4; these were items such as *relate–late*. For the moderate semantic set (Condition 3), the ratings were greater than or equal to 4 and less than 6 (e.g., *enlarge–large*), and in the high semantic set (Condition 4) the ratings were equal to or greater than 6 (e.g., *endanger–danger*).

To allow for the examination of phonological similarity in the absence of semantic similarity, and semantic similarity in the absence of phonological similarity, two additional conditions were created. Form only (Condition 1) consisted of 28 prime–target pairs that were phonologically transparent but semantically unrelated (e.g., *canine–nine*); these pairs were similar to those in the low semantic condition (e.g., *rehearse–hearse*), except that the form only items began with phonological segments that do not function as prefixes in any words; for example, the *ca–* in *canine* does not carry a systematic meaning that recurs in other words. The comparison between the low semantic and form only conditions provides evidence concerning effects of morphology independent of meaning. Significant facilitation for the low semantic condition (e.g., *relate–late*) but not for form only (e.g., *canine–nine*), would indicate that a factor other than semantic or phonological similarity (perhaps morphological structure) contributed to the priming effects because these conditions are equated with respect to semantic and phonological similarity. The absence of priming in both conditions would be consistent with the hypothesis that facilitation arises from semantic and phonological similarity rather than morphological relatedness.

Finally, in the semantic only set (Condition 5), the 28 word pairs were synonyms (most highly semantically similar) with no phonological similarity (e.g., *destiny–fate*). Sample stimuli for each condition and the mean similarity ratings are shown in Table 8.

For each of the 140 test primes (five conditions of 28 items each), a control prime was selected to match the test prime in frequency, number of syllables, and part of speech (see Table 9 for mean [Kučera & Francis, 1967] frequency values). Test and control primes were neither phonologically nor semantically related. In addition, to avoid experiment-specific response strategies, 140 varied nonword fillers were included, some phonologically related (e.g., *aspirin–pirin*) and others not (e.g., *eclipse–bort*). The phonologically related nonwords consisted of two types, words with pseudoaffixes (e.g., *recruit–cruit*) and words without recognizable affixes (e.g., *vibrate–brate*). This was done to ensure that participants could not develop a strategy whereby all phonologically related words or all words that ended in suffixes could be correctly responded to as real words. The items were divided initially into two lists, one with the test–target pair (e.g., *midnight–night*) and the other with the control–target pair (e.g., *wisdom–night*) so that each participant saw each target only once, preceded either by the corresponding test or control prime. Two separate, pseudorandom orders of all the items were generated to create a total of four lists. All of the test and control primes were digitally recorded by a female native English speaker.

Table 8
Sample Stimuli and Mean Similarity Ratings From Experiment 4 (Graded Semantic Similarity—Prefixes)

Condition	Prime–target example	Semantic similarity *M*	Semantic similarity *SD*
Form only	*coffee–fee*	1.50	0.8
Low semantic	*rehearse–hearse*	2.50	1.51
Moderate semantic	*midstream–stream*	5.10	2.07
High semantic	*preheat–heat*	7.30	1.34
Semantic only	*destiny–fate*	8.20	0.89

Note. 1 = *unrelated*, 9 = *highly related*.

Procedure

The procedure was exactly the same as described in Experiment 1. Eleven participants were tested on List 1, 11 on List 2, 10 on List 3, and 10 on List 4. The experiment took approximately 25 min to complete, including practice, warm-up, and test trials. Reaction times and lexical decision responses were recorded automatically by the computer.

Results and Discussion

Trials on which participants made an error (4.8%) and extreme values (responses greater than 2,000 ms or less than 200 ms: 1.0%) were excluded from the latency analysis. Decision latencies were entered into an ANOVA with the variables of prime type (test or control) and condition (the five types of prime–target relations: form only, low semantic, moderate semantic, high semantic, and semantic only). All means presented are based on participant analyses. Summary data are presented in Table 9. Mean decision latencies and error rates for the nonword items are presented in Appendix D.

The main effect of prime type was significant, $F(1, 41) = 12.2$, $MSE = 1710.00$, $pr\eta^2 = .23$. The main effect of condition was

Table 9
Mean Lexical Decision Latencies (in Milliseconds), Frequencies, and Error Rates for Experiment 4 (Graded Semantic Similarity—Prefixes)

		Reaction time		Control frequency			Reaction time		Test frequency			
Condition	Example	*M*	*SD*	*M*	*SD*	Error rate	*M*	*SD*	*M*	*SD*	Error rate	Priming effect
Form only	*coffee–fee*	616	144.10	17.50	5.40	.09	643	134.10	8.30	.13	18	−27*
Low semantic	*rehearse–hearse*	611	141.80	10.20	2.70	.07	602	160.30	4.50	.10	12.10	9
Moderate semantic	*midstream–stream*	569	140.50	9.20	2.60	.02	549	141.80	1.30	.01	4.60	20*
High semantic	*preheat–heat*	559	135.40	12.90	4.10	.02	517	137.50	5.70	.01	15.40	42*
Semantic only	*destiny–fate*	584	122.10	15.60	4.70	.02	557	140.30	5.50	.01	15.50	27*

* $p < .05$.

significant, $F(4, 164) = 51.5$, $MSE = 2247.00$, $pr\eta^2 = .56$. Although the stimuli were matched across conditions, decision latencies were shorter in the high semantic condition for targets following both test and control primes, indicating that these items may have been slightly easier overall. Finally, the Prime Type × Condition interaction was significant, $F(4, 164) = 6.1$, $MSE = 2342.00$, $pr\eta^2 = .13$. The differences between test and control primes in individual conditions are shown in Table 9.

In the form only (e.g., *canine–nine*) condition there was a significant interference effect of –27 ms, $F(1, 41) = 5.12$, $MSE = 3033.00$, $pr\eta^2 = .11$, whereas in the low semantic (e.g., *rehearse–hearse*) condition there was a nonsignificant effect of 9 ms, $F(1, 41) < 1$. The significant interference result for the form only condition is similar to the findings for the form only condition (*spinach–spin*) from Experiment 1 in which there was a trend toward interference that did not reach significance. These form only conditions differ from the low semantic ones in two important ways: (a) The form only test primes do not begin with pseudoprefixes, and (b) the pairs in the low semantic condition are slightly more semantically similar ($M = 2.5$ compared with 1.5). Items in the low semantic condition range in similarity up to 4 on the 9-point scale. These types of words produce more variable effects, with significant interference in some cases, nonsignificant, but numerically inhibitory effects in others, and slight, but nonsignificant facilitation in still others (cf. Allen & Badecker, 1999).

The moderate and highly semantically related conditions yielded significant priming effects, with the magnitude of the effect determined by the degree of similarity. Moderately related words (e.g., *midstream–stream*) primed about half as much (20 ms vs. 42 ms) as highly related words (e.g., *preheat–heat*). The effects in both of these conditions were significant: moderate semantic, $F(1, 41) = 4.29$, $MSE = 2005.00$, $pr\eta^2 = .09$; high semantic, $F(1, 41) = 29.0$, $MSE = 1270.00$, $pr\eta^2 = .41$. These results are remarkably similar to those found for suffixed–stem pairs in Experiment 1, in which highly related items (e.g., *teacher–teach*) produced a 40-ms facilitation, and moderately related items (e.g., *dresser–dress*) produced a significant effect of 19 ms (see Figure 2). The semantic only condition (pairs such as *destiny–fate*) also yielded a significant priming effect of 27 ms, $F(1, 41) = 5.7$, $MSE = 2630.00$, $pr\eta^2 = .12$.

It is important to note that a test of the 20-ms priming in the moderate condition compared with the 42-ms priming for the high condition showed that the high semantic items primed significantly more than the moderately similar pairs, $F(1, 41) = 4.44$, $MSE = 5851.00$, $pr\eta^2 = .10$, thus supporting the claim that morphological priming effects are related to semantic similarity and, even more crucially, that these effects are graded.

Data concerning trials on which participants made errors were also analyzed. The error rates were entered into an ANOVA with the variables of prime type (test or control) and condition (the five types of prime–target relations). Error rates by condition are shown in Table 9.

The effect of prime type was not significant, $F(1, 47) = 1.86$, $p = .18$, indicating that it was no more difficult to respond correctly to the targets when they were preceded by test primes compared with unrelated control primes. There was a significant main effect of condition, $F(4, 188) = 57.76$, because the form only and low semantic conditions produced more errors for targets following both the test and control primes than the other conditions. There was also a significant Prime Type × Condition interaction, $F(4, 188) = 5.65$. This result is similar to what we found for the form only condition in Experiment 1. The pattern of

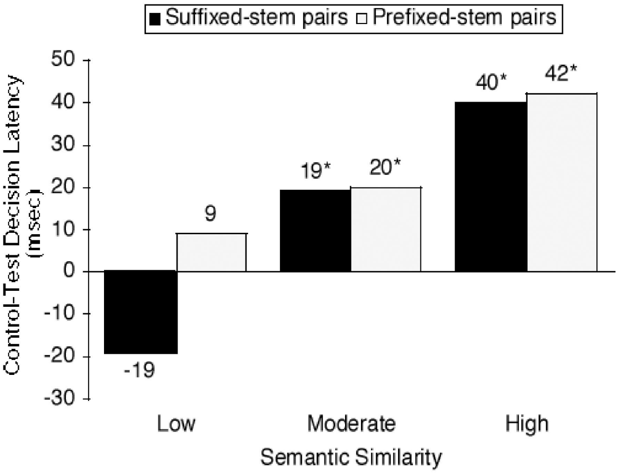

Figure 2. Comparison of the mean priming effects by condition for suffixed–stem and prefixed–stem pairs. Conditions vary in the degree of semantic similarity between primes and targets: Low (*hardly–hard* and *rehearse–hearse*); Moderate (*lately–late and midstream–stream*); and High (*boldly–bold* and *preheat–heat*). *$p < .05$.

results from prefixed words and stems provides further evidence that the magnitude of priming effects increases with increasing semantic similarity and that the effects are not restricted to suffixation.

GENERAL DISCUSSION

The principal goal of this research was to develop and test a theory of the bases of morphological structure and its role in processing complex words. The starting point for the research was the observation that although most linguistic and psycholinguistic theories assume that complex words consist of discrete morphemes, there is no consensus about the criteria for identifying morphological units, leaving the morphological status of many words unclear. Attention has tended to focus on clear cases that seem to consist of discrete morphemes, with less attention paid to structures that exhibit some but not all of the properties. Such cases present a challenge for the several existing theories that propose that complex words are processed in terms of component morphemes. It has been assumed that the principles identified by studying the clear cases will extend to the many words that exhibit partial regularities, but it is unclear how.

Our approach obviates this problem because it does not require partitioning words as morphologically simple versus complex. Rather, it involves identifying general principles concerning knowledge representation, learning, and processing that apply uniformly to all words. Taking the entire range of cases into consideration suggests an alternative to the standard approach: Morphemes are not like beads on a string but rather reflect degrees of convergence between form and meaning that vary across words (Seidenberg & Gonnerman, 2000). The cases that are puzzles for the classical view—such as the *–mit* and *cran–* structures—receive a natural interpretation within this theory: They are intermediate cases on a continuum ranging from strict morphological compositionality to monomorphemic. This view suggests that effects of morphological structure on processing should be predictable from the degree of semantic and phonological overlap between words and that these effects should be graded rather than categorical. The experiments addressed this claim as well as other phenomena previously thought to implicate discrete morphological units.

In this section, we summarize the main results and how they bear on the alternative theories. We also consider some findings that present challenges for a nondiscrete theory like ours and discuss several important unresolved issues that need to be addressed in further research.

Summary of Results

The main result of the experiments is that priming effects were predicted from semantic and phonological overlap rather than from morphological status (i.e., whether a word is typically considered to be morphologically complex). In Experiment 1, the critical stimuli all had the phonological form of morphologically complex words. With the phonological properties of the stimuli equated, the magnitudes of the priming effects were related to degree of semantic similarity. A word such as *boldly* behaves as though it contains the word *bold*, producing the largest priming effect. *Lately* bears the same phonological relationship to *late* but is less similar in meaning; this condition produces a smaller but significant priming effect. *Hardly* also overlaps with *hard* with respect to form, but the words are unrelated in meaning; this condition did not produce a significant priming effect. The magnitude of priming was significantly correlated with rated semantic similarity for all the suffixed–stem pairs. Thus there is a graded effect of semantic similarity across words that are phonologically similar. Such graded effects are consistent with the theory that morphological structures are not discrete units but rather reflect degrees of semantic and phonological similarity across words.

The first experiment also provided evidence concerning the role of formal overlap. The results replicated Marslen-Wilson et al.'s (1994) finding that words that are only related in form do not produce reliable priming. Marslen-Wilson et al. obtained this result with pairs such as *tinsel–tin*; we observed the same result with items such as *hardly–hard*. They concluded that formal overlap does not contribute to priming effects and so did not examine this factor in subsequent studies. However, other conditions in Experiment 1 show that this conclusion is too strong. The stimuli in the *teacher–teach* and *idea–notion* conditions were equated in terms of semantic similarity but differed in formal overlap; both conditions yielded significant priming, but the effect was larger for *teacher–teach*. Thus the effect of formal overlap depends on the degree of semantic similarity: Formal overlap only has an effect if prime and target are also semantically similar.

Experiment 2 provided further evidence concerning the effect of formal overlap. The critical stimuli were all highly semantically related. Given this degree of semantic similarity, there were graded effects of phonological structure: Words that were more phonologically similar produced larger effects. These results provide additional evidence against the claim that formal overlap does not contribute to priming effects for morphologically complex words; they also suggest that morphological structures are graded insofar as priming effects depended on the degree of formal overlap.

Experiment 3 addressed a claim that was central to Marslen-Wilson et al.'s (1994) morphological decomposition theory: that pairs of morphologically related words such as *observation–observant* do not produce priming. Marslen-Wilson et al. took this finding to indicate that the relations between such words are represented in lexical memory, but with inhibitory links between them. Our study suggests a different conclusion: As with the other conditions we have studied, such pairs do prime if they are sufficiently phonologically and semantically related (the *saintly–sainthood* condition). These stimuli were both phonologically transparent and highly semantically related, whereas the Marslen-Wilson et al. pairs were less similar to one another on both dimensions. The stimuli in our less phonologically related condition produced a 14-ms effect, very close in magnitude to that observed in the Marslen-Wilson et al. study (11 ms).

Experiment 3 also added further information about the interaction between semantic and phonological overlap. With a low level of semantic similarity, there was no effect of formal similarity (priming effect = −4 ms). With more highly semantically related pairs, pairs that exhibit greater phonological overlap (e.g., *saintly–sainthood*) produced a larger priming effect (34 ms) than did pairs that exhibit less phonological overlap (e.g., *observation–observant*; 14 ms). This result can only be taken as suggestive because the 14-ms effect in the *observation–observant* condition did not reach statistical significance. However, the pattern of results is consistent with those in the other experiments.

Results from Experiment 4 extended these results to a different type of derivational morphology, prefixed words, providing further support for the claim that semantic and phonological similarity strongly predict priming effects. The results from this experiment were strikingly similar to those found for the suffixed–stem word pairs in Experiment 1: Highly related prefixed words and stems (e.g., *preheat–heat*) primed twice as much (42 vs. 20 ms) as moderately related pairs (e.g., *midstream–stream*), and unrelated pairs (e.g., *rehearse–hearse*) yielded no priming. These results are inconsistent with previous claims that prefixed and suffixed words behave differently (e.g., Beauvillain, 1994; Colé, et al., 1989; Marslen-Wilson et al., 1994). This is not to deny that there are differences between prefixes and suffixes that affect the representation and use of lexical knowledge. For example, suffixes seem to be more prevalent in the world's languages (Cutler, Hawkins, & Gilligan, 1985). We also expect that there may be some subtle differences in auditory processing of prefixed and suffixed words on the basis of the time course of information availability. However, our results suggest that the internal structure of both types of words is determined by the convergence of semantic and phonological information.

In summary, the results of these experiments provide evidence consistent with a theory in which morphological structure is graded rather than discrete, reflecting the extent to which formal and semantic properties of words are correlated. The studies also show that several findings from the previous literature that had been taken as evidence for morphological decomposition can be understood in terms of our theory, specifically the importance of the degree of phonological and semantic similarity and the interaction between these two factors. The fact that a high degree of formal similarity is not sufficient to produce priming does not mean that formal similarity has no effect on processing; the effect depends on semantic similarity. Similarly, the fact that some suffixed–suffixed pairs do not prime does not reflect a generalization about this class of morphologically complex words; whether such pairs produce priming depends on how semantically and phonologically similar they are. Finally, to complete the circle, phonological similarity also exhibits graded effects, given that the words are sufficiently semantically related. Thus our approach subsumes many existing phenomena within a theory that also accounts for other data. This approach has the additional benefit of obviating the problem of finding a theoretical basis for establishing a boundary between morphologically simple versus complex words.

Methodological Issues

One concern about these studies is whether intermediate-sized priming effects seen in the critical moderately related conditions occurred because the stimuli were a mix of morphologically related and unrelated items. It could also be argued that the intermediate effects arise because some participants treat these words as morphologically complex and others treat them as morphologically simple. However, careful examination of the semantic similarity ratings shows that this is not the case. In fact, all of the participants were using the entire semantic rating scale, and the standard deviations were very low. This would not occur if the intermediate ratings arose from the averaging of bimodal distributions. Furthermore, histograms of the ratings for the intermediate items showed normal, not bimodal, distributions. These data suggest that we are correct in concluding that morphological priming reflects a continuum of similarity. Similarly, the intermediate effects were not due to a subset of morphemes (e.g., the "true" ones) because the various affixes appeared at each level of similarity (e.g., low: *hardly*; moderate: *lately*; high: *boldly*) and several affixes were used (e.g., *–able*, *–age*, *–ance*, *–er*, *–en*, *–ence*, *–ive*, *–ment*, and *-ly*). Finally, the objection that only some of the stimuli were truly morphologically complex would need to be based on independent criteria concerning morphological status, not a post hoc inspection of the data; as we have argued, there is little agreement about such criteria.

It will be necessary in future research to determine whether the same results obtain under other conditions. We used cross-modal priming with no delay between prime offset and target onset. Our results for immediate cross-modal lexical decision strongly support our approach, but the question remains as to how well they will generalize to other experimental paradigms. Priming for morphologically related words has also been observed intramodally (i.e., stimuli that are both visual or both auditory; Marslen-Wilson & Zhou, 1999). It will also be important to examine other stimulus onset asynchronies because semantic priming dissipates at longer lags (Feldman, 2000). It is unclear how the interactions between semantic and phonological similarity will change as the lag between stimuli increases. One possibility is that pairs that behave as though they are related at short stimulus onset asynchronies will behave as though they are unrelated at longer ones.

Empirical Challenges

Although our studies and the others to which the studies are closely related provide a consistent picture, it is important to consider several other findings that appear to fit less well. A prime example is the body of work on Hebrew, a language typologically distinct from English in that morphological structure is pervasive (Bentin & Feldman, 1990; Frost, Deutsch, & Forster, 2000; Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000; Frost, Forster, & Deutsch, 1997). For example, Bentin and Feldman (1990) found facilitation for morphologically related but semantically opaque words (e.g., *miktav* "letter" primes *katava* "article"). Such nonsemantic effects appear to challenge the distributed connectionist approach. If, as we argue, morphology arises from the conjunction of semantic and phonological codes, how do we deal with these nonsemantic effects? The answer lies in an additional property of the kind of connectionist approach we adopt, namely, the fact that all of the sound–meaning mappings must be learned within one system. The processing of all of the items, semantically related and unrelated forms alike, will be influenced by the degree to which the entire language system is morphologically structured. Because Hebrew is richer morphologically than English, priming may extend to semantically unrelated items. Plaut and Gonnerman (2000) carried out simulations in which a model was trained by using either a morphologically rich artificial language (like Hebrew) or an impoverished one (like English). They then tested priming for morphologically related word pairs that varied in their degree of semantic overlap. For both languages, the degree of semantic transparency affected the magnitude of priming, but semantically opaque items produced facilitation only in the morphologically rich language. The modeling results are consistent with the behavioral data and suggest that the model will pick up on the structure that is available in the input. Ubiquitous morphological structure coerces the model into adopting internal representations that more strongly preserve formal similarity, even when semantic representations diverge. Thus, data from typologically divergent languages not only are compatible with our account but also are better explained by the connectionist approach.

Some researchers have claimed to find nonsemantic (morphological) effects by using experimental paradigms that are considered insensitive to meaning. For example, in a visual masked priming experiment, Forster, Davis, Shoknecht, and Carter (1987) found that *cars* primes *car* but *card* does not. The primes are equated for orthographic overlap but differ both in morphological and semantic relatedness. The argument put forth by Frost et al. (1997) is that morphological, not semantic, relatedness must be causing the effects because pure semantic priming is typically not found in this paradigm (but see Bodner & Masson, 2003, for a counterexample). This kind of morphological masked priming occurs across several languages, including English (Forster et al., 1987), French (Grainger, Colé, & Segui, 1991), Dutch (Drews & Zwitserlood, 1995), and Hebrew (Frost et al., 1997). On first view, these effects would indeed seem to pose a problem for our approach. If masking prevents semantic priming, how can it be the semantic similarity, rather than the morphological relatedness, of *cars–car* that causes them to prime more than *card–car*? One possibility is that the effects of semantic similarity are only reduced by the mask but not completely eliminated; thus, visually dissimilar words will not prime (*gold–silver*), but visually similar ones may show differences in priming based on degree of semantic similarity. Gonnerman and Plaut (2000) presented results demonstrating that degree of semantic similarity does modulate masked priming effects; *boldly–bold* primes significantly more than *lately–late* or *hardly–hard*. Thus, the apparently nonsemantic effects from masked priming are actually influenced by semantic similarity.

Stolz and Besner (1998) also argued that their results are problematic for our account. They used a lexical decision task but required participants to perform a letter search first, which is supposed to remove the semantic component from the processing of the prime so that purely semantically related pairs (e.g., *gold–silver*) do not prime. In their experiments, *marked* primed *mark* more than did the orthographic control, *market*. However, recall that per our account form and meaning interact, so we would in fact predict that priming would be greater for items that are similar in both meaning and sound (e.g., morphologically related pairs) than for those related only on one of the two dimensions (form in *market–mark* or meaning in *gold–silver*). Indeed, the Stolz and Besner results, far from presenting a challenge, are entirely consistent with a distributed connectionist approach; word pairs in which semantic and phonological information converge produce greater priming (e.g., *marked–mark*), than those that diverge (e.g., *market–mark*). Our approach predicts that the degree of semantic similarity should also modulate effects in the letter search before priming paradigm, even if there is no pure semantic priming for orthographically unrelated items.

Theoretical Challenges

On our account, the similarity ratings reflect distance between semantically related items in a complex, high-dimensional semantic space. The exact dimensions composing that space are not easy to delineate. What does seem clear is that the ratings reflect similarity of global word meanings rather than the meanings of word parts. Thus, participants rated antonyms, such as *harmful–harmless*, relatively low in similarity, contrary to what one would predict if the judgments were based on stems, as *harm* has the same meaning in both forms. Word pairs differing in thematic role, such as *drinker* and *drinkable*, were also given lower ratings than might be expected if the ratings were based solely on the transparent stem *drink* being contained in both words. It is interesting that even these somewhat counterintuitive ratings are predictive of priming; the *harmful–harmless* and *drinker–drinkable* forms given lower ratings did not prime. Although the similarity ratings are clearly useful, further research is needed to explore exactly what the nature of semantic similarity is. Models of semantic similarity based on word co-occurrence statistics, for example the hyperspace analogue to language model (HAL), developed by Burgess and colleagues (Burgess, 1998; Burgess & Lund, 1997), and latent semantic analysis (LSA), developed by Landauer and colleagues (Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998), have made important contributions to our understanding of semantic similarity and its role in lexical processing. However, both of these models have difficulty approximating the similarity ratings produced by human raters, precisely in the domain of interest in this study, that is, derivationally related word pairs. Both models systematically underrate the similarity of derivationally related words, possibly because of the difference in grammatical category. Thus, *teacher* and *teach* are not as similar in HAL or LSA as they are in human ratings. HAL and LSA are much better able to predict human similarity judgments for inflectionally related pairs that do not cross grammatical categories, such as plurals or past tenses. In addition, these models systematically overrate the similarity of words that overlap in sound but not meaning, such as *spinach* and *spin.* Although models such as HAL and LSA have been extremely useful in explaining semantic similarity in many contexts, the predictive power of the models is weaker for precisely the cases of interest to us here, namely derivationally related words.

Another important aspect of the results reported here is that semantic similarity seems to play a more crucial role than phonological similarity, as pure semantic priming is found (e.g., *destiny–fate*), but pure phonological priming (e.g., *pigment–pig*) is not. Because semantic similarity has been shown to strongly affect processing, when examining phonological similarity in Experiment 2 we used word pairs that were highly semantically similar and very closely matched. Only when the pairs are highly related in meaning is it possible to see graded effects of phonological similarity. In English, then, phonological similarity between related words is a weak constraint, allowing its effects to be seen only at high levels of the stronger constraint, semantic similarity. This raises the question of why semantic similarity should be more important than phonological similarity. Although this is the case for a morphologically impoverished language such as English, the relative weighting of semantic and phonological influences on the system may be different for morphologically richer languages such as Hebrew. One explanation involves the nature of the space for each system. Phonological space is denser than semantic space. Moving slightly in phonological space, for example from *bank* to *bang*, causes a big change in meaning. However, moving a bit in semantic space, for example from a cat with a tail to a cat without a tail, does not have as large an effect. Semantics may also play a greater role than phonology because the semantic system is much richer. Semantic representations incorporate information from various input modalities, including visual, tactile, auditory, and kinesthetic domains, whereas phonological information is based on only one sensory modality. These differences between the semantic and phonological systems contribute to the primacy of semantics in processing complex words in English. In addition to varying cross-linguistically, the relative weighting of these constraints may vary within a language based on factors such as task parameters. For example, semantic similarity may not play as strong a role in a naming task as in a lexical decision task like ours or in a masked priming paradigm, in which primes are not consciously processed.

The Lexicon in the Distributed Connectionist Approach

In our view, representations of both form and meaning are distributed across many simple, neuron-like processing units. Comprehending a word involves computing its semantic representation from a phonological pattern, and producing a word involves computing the phonological pattern on the basis of an activated semantic pattern. Words are stored, not as independent, separate patterns but in the weights on connections between these simple processing units. All the words that a speaker knows are superimposed on the same set of weighted connections. However, the system should not be thought of as just a correlational system. There are certainly cases in which the correlations between sound and meaning are not strong, but in which there seem to be systematic relationships between words such as *–mit* in *submit*, *remit*, and *permit* (Aronoff, 1976). Such cases do not falsify the account because the intermediate representations that the system develops may be quite different from either the input or the output. The bank of processing units between the phonological and semantic units can transform the representations so that the relationship between the sound and meaning need not be simply correlated.

The differences between our account and more traditional accounts that reify morphology are subtle and concern the representational structure of morphological information. Both accounts accept that morphological effects exist, but they differ in terms of their explanation for the genesis of these effects. In the traditional account, morphological effects arise as the system decomposes complex words into stems and affixes in storage, access, or both. On our account, morphological effects arise as the consequence of the interactions in a dynamic system that maps meanings onto forms and vice versa. The system will learn to treat subcomponents of words componentially, to the extent that these components contribute systematically to the meaning–sound mapping. We do not stipulate the nature of morphemic structure, it arises naturally from independently motivated and characterized semantic and phonological systems.

Traditional discussions of morphology focused on parsimony issues. For example, it was argued that it would be more efficient to generate related forms by rule rather than to store each word separately. These arguments were based on intuition rather than on computational results; in fact, it is not clear whether the rule system needed to generate all and only the correct forms would be more or less complex than a system that listed each form separately. Questions about the trade-offs between storing and generating items do not arise in the type of system we have described because the network performs both functions. The units that encode knowledge of words and relations among them are also used in production and comprehension. Moreover, although such systems require a large number of units to encode a realistically sized lexicon (see, e.g., Harm & Seidenberg, 2004), economy is achieved because the same weights and units are used in processing all words. Thus, traditionalist arguments advocating morphological rules because they allow savings on otherwise high storage costs (e.g., Sandra, 1994) are less relevant in distributed connectionist systems.

In our system, then, a morpheme is simply a convenient description of a phonological segment that contributes systematically to the meanings of a set of words. Our notion differs from that classical definition of a morpheme as a minimal meaning-bearing unit in that it is graded; sometimes the same phonological segment contributes a lot to the meaning of a word (e.g., *–er* in *baker*), sometimes somewhat less (*–er* in *dresser*), and in some forms not at all (*–er* in *corner*). Morphology is a characterization of the structure present in the lexical system; it does not have independent status as a level of representation, and it is not possible to point to any place in the system where morphology is independently represented. There are not morphology-specific processing mechanisms; the same general principles that govern phonological and semantic processing of whole words and sentences govern the processing of the subparts of words commonly called morphemes.

The reason we have so strongly emphasized the graded nature of the priming effects found for both semantic and phonological similarity is that they provide the crucial evidence that can differentiate between a connectionist approach to morphology and more traditional, decomposition or dual-route theories. How would a theory that proposes storing stems and affixes separately determine which complex words to decompose? For example, in Marslen-Wilson et al.'s (1994) approach, semantically transparent forms are decomposed, but opaque forms are not. This theory is easy to apply to forms that are at either end of the semantic similarity continuum, but it is unclear how to handle intermediate forms, such as *lately–late*, within such an approach. Are intermediate forms decomposed or not? If they are not, then why do they prime one another? And if they are decomposed, then why do they prime less rather than more related forms? The same questions apply for other hybrid models such as the one proposed by Colé et al. (1989), who advocated decomposition for suffixed but not prefixed words. Two aspects of our results are particularly problematic for their model: First, the strikingly similar priming for both suffixed–stem and prefixed–stem word pairs, and, second, the graded priming effects for both affixation types. Finally, the results are also hard to accommodate in hybrid models that postulate a race between whole word mechanisms and decomposition, such as Frauenfelder and Schreuder's (1992) morphological race model. If the priming effects are graded, how does one determine where to draw the line to choose which mechanism processes which words? Thus, the intermediate effects that we have reported here both capture the generalizations one can make about morphology better than more traditional views and present a challenge to these views.

Of course, the type of model we are suggesting has yet to be implemented. We are making assumptions about the capacity of these models to discover the right sorts of generalizations that will allow them to learn the mapping from meaning to sound for a large body of words. There is evidence from models of word reading that these principles can account for a wide range of empirical findings (Harm & Seidenberg, 1999; Plaut et al., 1996). There is also evidence that these principles can account for some of the cross-linguistic differences in morphological processing (Plaut & Gonnerman, 2000). It remains to be seen whether large-scale models that use mulitsyllabic complex words can be developed and, if so, the extent to which they can account for the variety of empirical results in this domain.

It also will be important to flesh out the role syntactic information plays in representing and processing complex words. Derivational affixes often change the grammatical class of stems (e.g., adding *–er* changes the verb *bake* into the noun *baker*). Adding a suffix to a stem can also systematically change its syntactic properties. Words with the same morphemes, such as *runner*, *teacher*, and *baker*, also share distributional characteristics. This aspect of morphological structure raises important issues about the characterization of syntactic structure that we have not addressed. In order to account for such phenomena, the network we propose will need to also keep track of distributional and sequential information. How well such a network can account for syntactic phenomena is unknown.

CONCLUSIONS

To date, there has been little experimental evidence that can distinguish between connectionist approaches to lexical processing and more traditional approaches. The data we have presented are easily accounted for within a distributed connectionist approach, but are difficult to accommodate in all-or-none decomposition theories. On the distributed connectionist account, there are no decomposition or whole word procedures for lexical access. Per our theory, morphological regularities influence the development of interlevel representations that mediate mappings between semantics and phonology and that emerge in the service of language acquisition and processing. Morphology reflects structure present in the world: Language input contains patterns that are discovered by language learners to the extent that they are useful in solving the primary tasks of competent speakers, that is, comprehending and producing speech. Thus, although we assume these same principles operate across all languages, the system that emerges may differ depending on the reliability of phonological similarity as a cue to meaning, as well as other factors, such as the type and token frequencies of related complex forms and the nature of the orthographic system.

The present research represents a significant extension of concepts that were introduced in studies of monosyllabic, monomorphemic words to the processing of morphologically complex words. The results exhibit signature effects—such as the graded, interactive effects of semantic and phonological similiarity in producing morphological priming effects—that seem more compatible with models based on connectionist principles than standard models of the lexicon. Our account obviates the distinction between whole word processing and lexical decomposition as well as the traditional concept of discrete morphemes. Many important issues remain to be addressed. There are additional morphological phenomena in English to consider (e.g., compounding), as well as morphological phenomena that occur in other languages (e.g., reduplication), questions as to whether the same approach can be extended to languages that differ typologically from English and make even heavier use of derivational morphology (e.g., Hebrew), how such systems are learned, the breakdowns in morphological knowledge that occur as a consequence of some types of brain injury, and whether implemented connectionist models trained on realistic corpora will pick up on the same regularities as people apparently do. It also remains to be determined whether other types of models can account for both the data we have presented and these additional phenomena. Derivational morphology is thus likely to be a major focus of future research on language acquisition and processing and their brain bases.

References

Allen, M., & Badecker, W. (1999). Stem homograph and stem allomorphy: Representing and processing inflected forms in a multilevel lexical system. *Journal of Memory and Language, 41,* 105–123.

Andersen, E. S. (1992). Complexity and language acquisition: Influences on the development of morphological systems in children. In J. A. Hawkins & M. Gell-Mann (Eds.), *The evolution of human languages,* (pp. 241–271). Redwood City, CA: Addison-Wesley.

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

Aronoff, M. (1976). *Word formation in generative grammar.* Cambridge, MA: MIT Press.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173–218). New York: Cambridge University Press.

Bates, E., & MacWhinney, B. (1987). Competition, variation and language learning. In B. MacWhinney (Ed.), *Mechanisms of language learning* (pp. 157–193). Hillsdale, NJ: Erlbaum.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–73). New York: Cambridge University Press.

Beauvillain, C. (1994). Morphological structure in visual word recognition: Evidence from prefixed and suffixed words. *Language and Cognitive Processes 9,* 317–339.

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

Bodner, G. E., & Masson, M. E. (2003). Beyond spreading activation: An influence of relatedness proportion on masked semantic priming. *Psychonomic Bulletin and Review, 10,* 645–652.

Bradley, D. (1980). Lexical representation of derivational relation. In M. Aronoff & M.-L. Kean (Eds.), *Juncture* (pp. 37–55). Saratoga, CA: Amma Libri.

Buchner, A., Faul, F., & Erdfelder, E. (1992). GPOWER: A priori-, post hoc-, and compromise power analyses for the Macintosh [Computer program]. Bonn, Germany: Bonn University.

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

Burani, C., Salmaso, D., & Caramazza, A. (1984). Morphological structure and lexical access. *Visible Language, 18,* 419–427.

Burgess, C. (1998). From simple associations to the building block of language: Modeling meaning in memory with the HAL model. *Behavior Research Methods, Instruments, and Computers, 30,* 188–198.

Burgess, C., & Lund, K. (1997). Modeling parsing constraints with high-dimensional context space. *Language and Cognitive Processes, 12,* 177–210.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production* (pp. 257–294). San Diego, CA: Academic Press.

Bybee, J. (1985). *Morphology: A study of the relation between meaning and form.* Philadelphia: Benjamins.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes, 10,* 425–455.

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

Cassidy, K. W., Kelly, M. H., & Sharoni, L. J. (1999). Inferring gender from name phonology. *Journal of Experimental Psychology: General, 128,* 362–381.

Chomsky, N. (1970). Remarks on nominalization. In R. A. Jacobs & P. S. Rosenbaum (Eds.), *Readings in English transformational grammar,* (pp. 184–221). Waltham, MA: Ginn.

Christiansen, M., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes, 13,* 221–268.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers, 25,* 257–271.

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

Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Cutler, A., Hawkins, J. A., & Gilligan, G. (1985). The suffixing preference: A processing explanation. *Linguistics, 23,* 723–758.

Drews, E., & Zwitserlood, P. (1995). Morphological and orthographic similarity in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1098–1116.

Erdfelder, E., Faul, F., & Buchner, A. (1996). G*Power: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28,* 1–11.

Feldman, L. B. (1995). *Morphological aspects of language processing.* Hillsdale, NJ: Erlbaum.

Feldman, L. B. (2000). Are morphological effects distinguishable from the effects of shared meaning and shared form? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1431–1444.

Forster, K. I., & Chambers, S. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627–635.

Forster, K. I., Davis, C., Shoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 39*(A), 1098–1116.

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

Frauenfelder, U. H., & Schreuder, R. (1992). Constraining psychological models of morphological processing and representation: The role of productivity. In G. E. Booij & J. van Marle (Eds.), *Yearbook of morphology 1991* (pp. 165–183). Dordrecht: Kluwer Academic.

Frost, R., Deutsch, A., & Forster, K. I. (2000). Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 751–765.

Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., & Marslen-Wilson, W. (2000). Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory & Cognition, 28,* 1277–1288.

Frost, R., Forster, K., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 829–856.

Gonnerman, L., & Plaut, D. C. (2000). Semantic and morphological effects in masked priming. *Journal of Cognitive Neuroscience, 49E,* 116.

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

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106,* 491–528.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111,* 662–720.

Henderson, L. (1985). Toward a psychology of morphemes. In A. W. Ellis (Ed.) *Progress in the psychology of language* (Vol. 1, pp. 15–71). Hillsdale, NJ: Erlbaum.

Hockett, C. F. (1958). *A course in modern linguistics.* New York: Macmillan.

Jarvella, R. J., & Snodgrass, J. G. (1974). Seeing ring in rang and retain in retention: On recognizing stem morphemes in printed words. *Journal of Verbal Learning and Verbal Behavior, 13,* 590–598.

Joanisse, M., & Seidenberg, M. S. (1999). Impairments in verb morphology following brain injury: A connectionist model. *Proceedings of the National Academy of Sciences, 96,* 7592–7597.

Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99,* 349–364.

Kempley, S. T., & Morton, J. (1982). The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology, 73,* 441–454.

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

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211–240.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25,* 259–284.

Li, P., & MacWhinney, B. (1996). Cryptotype, overgeneralization, and competition: A connectionist model of the learning of English reversive prefixes. *Connection Science: Journal of Neural Computing, Artificial Intelligence & Cognitive Research, 8,* 3–30.

MacWhinney, B., & Leinbach, A. J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition, 40,* 121–157.

Manelis, L., & Tharp, D. A. (1977). The processing of affixed words. *Memory & Cognition, 5,* 690–695.

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

Marslen-Wilson, W., & Zhou, X. (1999). Abstractness, allomorphy, and lexical architecture. *Language and Cognitive Processes, 14,* 321–352.

Meunier, F. & Segui, J. (1999). Frequency effects in auditory word recognition: The case of suffixed words. *Journal of Memory and Language, 41,* 327–344.

Mirkovic, J., MacDonald, M. C., & Seidenberg, M. S. (2005). Where does gender come from? Evidence from a complex inflectional system. *Language and Cognitive Processes, 20,* 139–168.

Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the Logogen model. In P. A. Kolers, M. Wrolstad, & H. Bouma (Eds.), *Processing of visual language* (pp. 259–268). New York: Plenum Press.

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

Napps, S. E. (1989). Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory & Cognition, 17,* 729–739.

Napps, S. E., & Fowler, C. A. (1987). Formal relationships among words and the organization of the mental lexicon. *Journal of Psycholinguistic Research, 16,* 257–272.

Plaut, D. C., & Gonnerman, L. M. (2000). Are nonsemantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes, 15,* 445–485.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in a quasi-regular domain. *Psychological Review, 103,* 56–115.

Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multilayered perceptron: Implications for child language acquisition. *Cognition, 38,* 43–102.

Plunkett, K., & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition, 48,* 1–49.

Rayner, K. & Pollatsek, A. (1989). *The psychology of reading.* Upper Saddle River, NJ: Prentice-Hall.

Rueckl, J., Mikolinski, M., Raveh, M., Miner, C., & Mars, F. (1997). Morphological priming, fragment completion, and connectionist networks. *Journal of Memory and Language, 36,* 382–405.

Rumelhart, D., & McClelland, J. (1986). On learning the past tenses of English verbs. In J. McClelland, D. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 2: Psychological and biological models* (pp. 216–271). Cambridge, MA: MIT Press.

Sandra, D. (1994). The morphology of the mental lexicon: Internal word structure viewed from a psycholinguistic perspective. *Language and Cognitive Processes, 9,* 227–269.

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention and performance XII: Reading* (pp. 245–263). Hillsdale, NJ: Erlbaum.

Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences, 4,* 353–361.

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

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

Stolz, J. A., & 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.

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

Taft, M., & Forster, K. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14,* 638–647.

Appendix A

Experiment 1: Mean Lexical Decision Latencies and Error Rates for Nonwords by Condition (Graded Semantic Similarity—Suffixes)

Condition	Prime–Target example	Mean reaction time (ms)	Mean error rate
Phonologically related, change in target	*computation–compuse*	751	.03
Phonologically related, no “suffix”	*bishop–bish*	752	.07
Phonologically related, pseudosuffix	*slither–slith*	767	.09
Phonologically unrelated, suffixed prime	*hostess–dight*	747	.05
Phonologically unrelated, no suffix	*basil–grook*	771	.06

Appendix B

Experiment 2: Mean Lexical Decision Latencies and Error Rates for Nonwords by Condition (Graded Phonological Similarity)

Condition	Prime–Target example	Mean reaction time (ms)	Mean error rate
Phonologically related	*territory–territ*	833	.07
Phonologically unrelated	*boomerang–jaulic*	795	.04

Appendix C

Experiment 3: Mean Lexical Decision Latencies and Error Rates for Nonwords by Condition (Suffixed–Suffixed pairs)

Condition	Prime–Target example	Mean reaction time (ms)	Mean error rate
Phonologically related, pseudo suffix	*respectful–respection*	840	.12
Phonologically related, no “suffix”	*stylist–styleem*	738	.02
Phonologically unrelated, pseudosuffix	*optimal–brovian*	764	.06

Appendix D

Experiment 4: Mean Lexical Decision Latencies and Error Rates for Nonwords by Condition (Graded Semantic Similarity—Prefixes)

Condition	Prime–Target example	Mean reaction time (ms)	Mean error rate
Phonologically related, no “prefix”	*berserk–serk*	729	.04
Phonologically related, pseudoprefix	*inert–ert*	737	.05
Phonologically unrelated, prefixed prime	*dislike–preet*	702	.03
Phonologically unrelated, no prefix	*lotion–bleam*	723	.05

Received October 22, 2002
Revision received September 5, 2006
Accepted September 12, 2006 ■