SCIENTIFIC STUDIES OF READING, *7*(3), 219–237

Morphology in Word Identification: A Word-Experience Model That Accounts for Morpheme Frequency Effects

Erik D. Reichle and Charles A. Perfetti

University of Pittsburgh

In reading research, morphological processing and monomorphemic word identification have generally been treated separately. We describe a computational model that brings both kinds of reading together within a single framework. This model assumes that word knowledge—the orthography, phonology, and meaning of words—accumulates with experiences with individual words and that this knowledge is reflected in two functionally different aspects of word processing—*familiarity* and *availability*. We report simulations that demonstrate that the model accounts both for classical effects of frequency and consistency in simple word reading and for morphological effects in the reading of complex words. The morphology simulations naturally capture a distinction between inflectional and derivational morphology without defining this distinction a priori. We discuss the implications of our model for general issues in reading, including individual differences in reading ability.

Word identification models have largely ignored the role of morphology in lexical processing. In this article we demonstrate that a general model of word identification—that is, one that can account for a variety of simple phenomena in word identification—can also handle phenomena related to morphology. In particular, we demonstrate that the traditional distinction between derivational and inflectional morphology and their associated frequency effects can be explained using a single set of computational principles.

Requests for reprints should be sent to Erik D. Reichle, Department of Psychology, 635 Learning Research & Development Center, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15260. E-mail: reichle@pitt.edu

Theories of morpheme processing can be classified according to how they explain the identification of polymorphemic words. Proponents of the *compositional* accounts claim that the meanings of complex words are constructed from their component morphemes (Jarvella & Meijers, 1983; MacKay, 1978). From this perspective, a complex word is first parsed into its components so that these individual units of meaning can be identified and then assembled into its overall meaning. For instance, the meaning of *cowboys* would be ascertained by first parsing the word into its components and then assembling its meaning from these components (e.g., *cow* + *boy* + *s*).

In contrast, proponents of the *full-listing* theories maintain that complex words are represented in their entirety (Bradley, 1980; Butterworth, 1983; Henderson, Wallis, & Knight, 1994; Kempley & Morton, 1982; Manelis & Tharp, 1977; Rubin, Becker, & Freeman, 1979). By this view, *cowboys* would be represented as a single entity, with separate representations for words like *cow* and *boy*, and even for the singular form of the word, *cowboy*. These two views are not mutually exclusive, and theories that include both separated and compositional representations have been proposed (Caramazza, Laudanna, & Romani, 1988; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Niemi, Laine, & Tuominen, 1994; Taft, 1994). In these hybrid theories, complex words are identified via a "race" between compositional and whole-word lookup processes. These theories highlight that some words are more amenable to decomposition than others.

In this article we focus on a distinction related to this decomposition question: the distinction between *inflectional* and *derivational* morphology. Inflected morphology is generated through syntactic affixes that systematically control grammatical agreement while preserving the core meaning of the base form (the *stem*). Thus, the English inflectional system adds *s* to form a plural of a noun (*dog* → *dogs*) and to form the third-person singular of a verb (*run* → *runs*). The inflectional system reflects a closed rule-based process, as demonstrated by the fact that the plural forms of new nouns are readily generated; for example, the plural of *tark* is *tarks*. Whether the psycholinguistic process that implements this formal system also uses rules (Prasada & Pinker, 1993) or merely generalizes based on input "similarity" (Hare, Elman, & Daugherty, 1995; Rumelhart & McClelland, 1986) is a hotly disputed issue. However, people so consistently use the regular inflectional rule as the default in the face of competing similarity factors (Berent, Pinker, & Shimron, 1999) that this is a compelling reason to think that the human language process uses a rule system to generate inflectional morphology.

In contrast to inflectional morphology, derivational morphology works across grammatical categories, generating forms of a base morpheme in different grammatical categories. For example, *beauty, beautiful*, and *beautify* show a derivational pattern from a single base morpheme. Creativity in derivational morphology is possible to an extent not tolerated by the grammatical system in inflectional morphology. Note, for example, that the derivational paradigm for *beauty* may also include *beaut, beauteous,* and *beautician*. In addition to its greater (but still constrained) creativity, the derivational system can produce a marked change in the underlying meanings of base forms. Whereas the inflectional system creates only those meaning changes narrowly defined within the grammar (e.g., plural inflection means more than one), the derivational meaning changes are partly systematic—*beauty* and *beautiful* relate in the same ways as *plenty* and *plentiful*—and partly not. For example, *creative* and *creation* do not seem to be related in quite the same ways as are *secretive* and *secret*. And for a new noun (e.g., *tark*), generating an adjectival form is a bit less obvious (*tarkal? tarkish?*).

More generally, derivational forms are subject to variations in meaning transparency—the degree to which the meaning of a complex word is inferable from its components. For some words, the meanings of complex forms are *transparent*; in others, the meanings are *opaque*. In some cases, an opaque relation can be seen once attention is drawn to it. For example, many readers will not be aware at first that *discord* is related to *concordance* but will see the relationship when the two words are shown together. Others are related only through a language history that is unknown to ordinary users and remains opaque without serious study of Latin, Greek, and early French and Germanic languages. For example, *disdain* relates to *dignity* (*disdain* is the withholding of dignity) through the old French variant, *deignie*r (from the Latin *dignus*) and its derivational partner, *desdegneir.* But who knew? The degree of transparency partially determines how readily one might extrapolate the rules that govern the morphemic composition of complex words and the likelihood that any compositional process can be used (Marslen-Wilson et al., 1994).

We cannot address all of the many issues in morphology within our general model. Instead, we focus on two very general and important variables that will explain aspects of morphology (identifying complex words) and traditional (monomorphemic) word identification simultaneously. We want a single model that is general across simple and morphologically complex words, not one for which addition mechanisms are postulated for morphologically complex words. The advantages of simplicity and generality are obvious as a starting point, and it is conceivable that complexities will be forced eventually. Meanwhile, we have developed a model that can account for (a) frequency of occurrence and (b) the similarity among the items (words and/or morphemes) in a highly general manner. We focus on these variables because their effects have been well-documented and extensively studied in both the standard word-identification (Balota & Chumbley, 1984, 1985; Schilling, Rayner, & Chumbley, 1998; Seidenberg, 1985; Seidenberg, Waters, Barnes, & Tanenhaus, 1984) and morphology (Bertram, Baayen, & Schreuder, 2000; Hyönä & Pollatsek, 1998; Pollatsek, Hyönä, & Bertram, 2000; Schreuder & Baayen, 1997) literatures.

AN EPISODIC THEORY OF WORD IDENTIFICATION

Our central idea is that the processes that allow words meanings, pronunciations, and spellings to be encoded, stored, and retrieved in episodic memory may be sufficient to support word identification and may also explain how the morphological composition of words affects this process. Our starting assumption is that the mechanisms of episodic memory, in conjunction with the statistical properties of the information being represented (i.e., a reader's vocabulary), is sufficient to account for the myriad effects that both linguistic (e.g., frequency of occurrence) and perceptual (e.g., word length) variables have on the relative ease of using words. This dynamic view of lexical processing is contrary to the abstractionist view that words are represented by static representations (Tenpenny, 1995). Because this article focuses on morphology, our discussion is limited to several well-documented linguistic variables. A more complete discussion of the model and its theoretical coverage is forthcoming (Reichle, Landi, & Perfetti, 2002).

MINVERVA 2

Our model is implemented within the framework of an existing model of episodic memory: MINERVA 2 (Hintzman, 1984). This model simulates a variety of memory processes (e.g., learning, categorization) and is amenable to our conceptualizations of word identification and the manner whereby readers acquire this skill. Like other *instance-based* models (Logan, 1988; Medin & Schaffer, 1978; Nosofsky, 1992), the key assumption in MINERVA 2 is that episodic memory consists of memory traces and the operations that encode and retrieve them.

Individual memory traces represent specific experiences (e.g., an encounter with a new word) and include both *focal* information (e.g., the word's spelling and meaning) and *contextual* information (e.g., the physical setting in which the word is encountered). This information is represented as perceptual and/or semantic features. Thus, each experience with the word provides an opportunity to encode its spelling, pronunciation, and/or context-specific meaning. This happens whenever those features of the word being attended to (in working memory) are encoded as a new trace in long-term memory. Over one's lifetime, these experiences produce a cornucopia of such traces and a rich working knowledge of the words in one's language.

The information in long-term memory can be used in two ways. The first corresponds to the process of recognition. If a word that has been experienced before is encountered again, then the person may recognize the word as having been previously experienced (i.e., it is familiar). In MINERVA 2, familiarity is used to simulate recognition (Hintzman, 1987, 1988). In our simulations, *familiarity* is an index of how well a word is established in a reader's vocabulary—words experienced often tend to be more familiar than new words or words experienced less often.

The second way information in long-term memory can be used is through the retrieval of features. For example, one might learn that a *caracal* (pronounced \ˈkar-ə-ˌkal\) is a medium-sized cat indigenous to Africa and parts of the Middle East. Having learned this, if one sees the word again, it may be possible to both say the word and describe what it refers to. This capacity depends on the ability to use the word's orthographic form as a cue to retrieve its pronunciation and meaning. In MINERVA 2, this capacity is used to simulate recall (Hintzman, 1986). In our simulations, it is used to simulate the retrieval of the orthographic, phonological, and/or semantic features that compose words. The quality of this information and the facility with which it is retrieved is quantified using the Pearson's correlation coefficient, r, between the retrieved features and the features that (in our simulations) define a given word; this index provides a measure of a word's *availability*.

Additional Assumptions

Each experience with a word is encoded as a memory trace partitioned into three separate feature fields, which represent the word's orthography, phonology, and meaning. Similarity was instantiated by varying the amount of overlap among these features. For example, the memory traces for the homophones *bear* and *bare* contain partially overlapping orthographic features, identical phonological features, and completely orthogonal semantic features. Although our definition of "similarity" is completely arbitrary, the simulation results are robust and only depend on the relative amounts of feature overlap (e.g., homophones share more phonological features than nonhomophones).

Most existing models of word identification have been evaluated by teaching the model small corpora of words (e.g., 2,987 words in Seidenberg & McClelland, 1989) and then examining how well the model can use this knowledge to perform simulated tasks (e.g., pronunciation). In some cases this training is necessary to address theoretical questions (Patterson, Seidenberg, & McClelland, 1989), but most often it only intends to show that the model can retain information about a large number of words. These demonstrations are necessary (in connectionist models) because the benefits of early training are often eliminated or attenuated by subsequent training (Ratcliff, 1990).

With instance-based models, these precautions are unnecessary because individual experiences are represented by separate memory traces. Our model is thus capable of encoding, retaining, and retrieving any number of individual words. In fact, it was necessary to degrade our model's performance by adding a "noisy" memory trace representing the cumulative interference that would result from a lifetime's worth of experiences (see Appendix).

Another reason for not using a large corpus is that it is prohibitively time-consuming, both in terms of generating the corpus (which, if it is to be representative, must be much larger than existing corpora and contain complex, polymorphemic words) and in terms of running the simulations. Consequently, we opted to run a series of "test case" simulations involving only a few words having very specific, perfectly controlled characteristics. These simulations are existence proofs that the assumptions of the model are sufficient to explain certain phenomena (e.g., word frequency effects).

Our final assumption relates to the tasks that we simulated. Rather than attempting to simulate the mechanics of the various tasks that have been used to study word identification (e.g., pronunciation, lexical decision), we simply compared the model's relative performance across experimental conditions using the two dependent measures: *familiarity* and *availability*. To review, familiarity is an index of how well a particular word is known, whereas availability (as measured by *r*) refers to the quality of the word's identity and the facility with which it is retrieved from memory. Because availability captures both pronunciation and meaning, it is a composite measure of word identity.

SIMULATION RESULTS

Our first goal was to demonstrate that the model handles some general, important facts about word identification—specifically, word frequency and form similarity. For this, our simulations use monomorphemic words. Our second goal was to extend the model to morphologically complex words, showing that the same basic assumptions work for complex words.

Word Identification

The first set of simulations demonstrated the model's ability to capture two important word effects that are well established in the experimental literature.

Word frequency effects. The first simulation[1] examined how variation in frequency of occurrence affects a word's familiarity and the availability of its pronunciation and meaning. Six "words" covering the full range of frequency values (as tabulated by Francis & Kučera, 1982) were first encoded into the model's "vocabulary"; then the orthographic features of each word were successively presented to the model (i.e., used to probe memory) to determine the word's familiarity, pronunciation, and meaning. The results of this simulation are presented in Panel A of Figure 1, which shows the words' *availability* and *familiarity* as a function of their frequency.

[1]Each Monte Carlo simulation reported in this article is based on 1,000 statistical subjects per condition.

Panel A of Figure 1 shows that both familiarity and availability increase with word frequency. These results are consistent with frequency effects in a variety of tasks, including lexical decision (Balota & Chumbley, 1984) and pronunciation (Balota & Chumbley, 1985; Frederiksen & Kroll, 1976; Stanovich & West, 1981, 1983). To account for performance in these tasks using our model, it is necessary to assume only that lexical decisions can be made using a word's familiarity or identity (pronunciation and meaning), whereas reading aloud (and possibly other tasks; e.g., categorization) depends on a word's identity (i.e., its pronunciation and meaning).

Our distinction between word familiarity and identity also provides a theoretical bridge to models that link word identification to eye-movement control during reading (Engbert & Kliegl, 2001; Engbert, Longtin, & Kliegl, 2002; Kliegl & Engbert, in press; Reichle, Pollatsek, Fisher, & Rayner,1998; Reichle, Rayner, & Pollatsek, 1999, in press); namely, the distinction between a rapid familiarity assessment (which indicates that access to the meaning of the word in question is im-

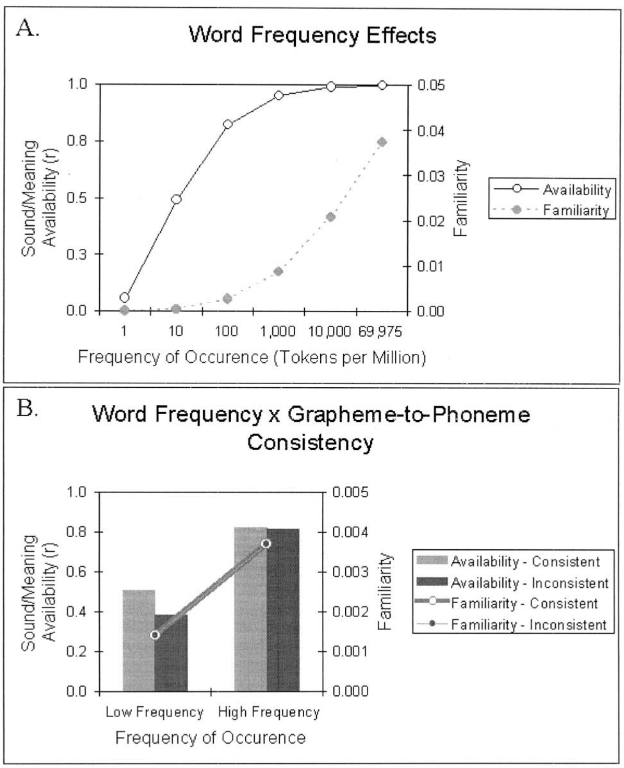

FIGURE 1 Panel A: Simulated word frequency effects. Panel B: Simulated interaction between word frequency and grapheme–phoneme consistency.

minent) followed by the process of actually identifying the word. This is an important step forward because it provides a basis for our model to explain word frequency effects in natural, silent reading (Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Schilling et al., 1998). It also raises the prospect that we might eventually be able to evaluate predictions of our model using eye-tracking methods in the context of natural reading experiments.

Consistency effects and their interaction with word frequency. The effects of word frequency interact with the word's spelling–pronunciation consistency. The typical finding is that low-frequency words having inconsistent correspondences are processed less rapidly than low-frequency words having consistent correspondences (Andrews, 1982; Seidenberg, 1985; Taraban & McClelland, 1987; Waters & Seidenberg, 1985). For example, it generally takes longer to say the word *pint*, which is pronounced differently than other words ending in -*int* (e.g., *mint*, *hint*), that it does to say *book*, which has a pronunciation consistent with other words ending in -*ook* (e.g., *look*, *cook*). This consistency effect is weaker or absent with high-frequency words; frequent inconsistent words are named as rapidly as frequent consistent words[2].

We predicted that our model would simulate the interaction between frequency and consistency because the information about a word reflects the global contents of memory and is a weighted function of both the frequency with which it is represented and its similarity to other information in memory (see Appendix). We therefore completed a second simulation in which we examined the model's performance in four conditions corresponding to the factorial manipulation of a target's frequency and similarity to other, nontarget items. In all of these conditions, one target and five nontargets words were encoded. The semantic features of all six words were completely orthogonal, but their orthographic features were similar. On average, the words shared 85% of the same orthographic features. In the two consistent conditions, the phonological features of all of the words were similar (i.e., 85% feature overlap), whereas in the two inconsistent conditions, the phonological features of the target words were only moderately similar to those of the nontargets (i.e., 50% feature overlap). Thus, the mappings between specific orthographic and phonological features were entirely consistent across the targets and nontargets in the consistent conditions but consistent only among the nontargets (and not between the targets and nontargets) in the inconsistent conditions. Finally, in two of the conditions, the target words were low frequency (10 per mil-

[2]Although small consistency effects can be observed among high-frequency words (Jared, 1997), the basic interaction is quite robust over studies using pronunciation (Andrews, 1982; Seidenberg, 1985; Taraban & McClelland, 1987; Waters & Seidenberg, 1985), lexical decision (Sereno, Rayner, & Posner, 1998), and reading (Sereno et al., 1998).

lion), whereas in the other two conditions, the target words were high frequency (100 per million).

The simulation results are presented in Panel B of Figure 1. The horizontal bars indicate that the model predicted an interaction between frequency and consistency: Word identities (pronunciation and meanings) of low-frequency inconsistent words are less available than those of both low-frequency consistent words and high-frequency words. It is also interesting that the model does not predict this interaction for word familiarity, which is a simple function of word frequency. This suggests that the interaction between frequency and consistency may be absent—or at least attenuated—in tasks that can be performed on the basis of familiarity alone (e.g., speeded lexical decision).

On the basis of these results, we conclude that the model promises to be a viable alternative to existing word identification models. Indeed, an earlier version of the model has already been used to explain the outcome of an experiment involving both word form (orthographic and phonological) and semantic priming (Reichle & Perfetti, 2001). Although these results will not be discussed here, the model handles such effects by instantiating the *compound-cue* notion of priming (McKoon & Ratcliff, 1992; Ratcliff & McKoon, 1988); to the extent that two words share overlapping features, the features of one word can provide a useful retrieval cue for facilitating the processing of the other.

Summary. The preceding simulations demonstrate the model's capacity to account for several basic phenomena in the word-identification literature. In the next section of this article, we demonstrate how the model accounts for several nonobvious effects related to morpheme frequency and similarity.

Morpheme Effects

The simulations reported here address questions about morphology related to frequency, meaning, and compositionality. For example, if encounters with morphological variants of a word build up the representation of its base form, then the model's ability to capture this would simulate token frequency effects (Bertram et al., 2002; Schreuder & Baayen, 1997).

Effects of token frequency of the inflected forms. The first simulation examined the effect of the token frequency of a base word's inflected forms on the base word itself. For example, do experiences with *run, running,* and *runs* add to the representational strength (familiarity or availability) of the stem *run*? Our simulation involved two conditions: One in which a base word had a low-frequency (10 per million) inflection, and a second in which a base word had a high-frequency (100 per million) inflection. The frequency of the base word was held constant across the conditions (10 per million), as was the similarity between the base words and their inflected form, which shared 85% of their orthographic, phonological, and semantic features. The orthographic features of the base words were used as probes to determine their familiarity and the availability of their pronunciations and meanings (i.e., their identities). The results of this simulation are presented in Panel A of Figure 2.

Panel A of Figure 2 shows that both the familiarity and the availability of the stems are affected by the token frequencies of their inflected forms. The base word with a high-frequency inflection is more familiar than the base word with an infrequent inflection, and the pronunciation and meaning of the high-inflection base

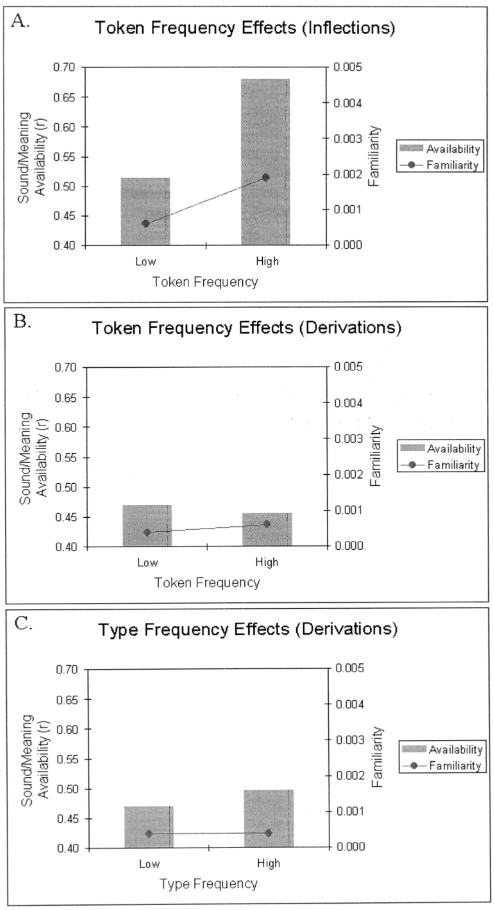

FIGURE 2 Panel A: Simulated effects of inflectional token frequency. Panel B: Simulated effects of derivational token frequency. Panel C: Simulated effects of derivational type frequency.

word are more available than those of the low-inflection base word. This pattern has been reported in the experimental literature (Bertram et al., 2000; Schreuder & Baayen, 1997).

Effects of token frequencies of derived forms. The second simulation asks whether the base forms of words are affected by the token frequencies of their derivations. This simulation was identical to the previous one, except that the similarity between the base forms and their derivations was reduced because, on average, derived forms are less orthographically, phonologically, and semantically similar to their base words than are inflected forms. Thus, the feature overlap between the base words and their derivations was reduced to 50%. The base words' orthographic features were then used to probe memory to again determine their familiarity and availability.

The simulation results in Panel B of Figure 2 indicate that neither the familiarity of the base words nor the availability of their pronunciations and meanings was affected by the token frequencies of their derived forms. The base word with the infrequent derivation was just as familiar as the base word with the frequent derivation, and the two base words' pronunciations and meanings were equally available. Again, this pattern has been reported in the experimental literature (Bertram et al., 2000; Schreuder & Baayen, 1997).

Effects of derivational type frequency. Our final simulation addressed a paradoxical finding that the type frequency, or *family size* (Schreuder & Baayen, 1997), of derivational forms—the number of different derivations that a word has, rather than the token frequency across these derivations—affects processing of the base word. Thus, words with many derivational forms (e.g., *observe*: *observer*, *observation*, *observant*, *observance*, *observable*, *observatory*) are typically identified more rapidly than words with few derivational forms (e.g., *poison*: *poisonous*), even when the token frequency across the forms is controlled. This simulation thus contrasted two conditions. In the first, a base word (frequency = 10 per million) and *one* derivation (frequency = 10 per million) were encoded. As in the previous simulation, the derivation shared 50% feature overlap with its base form. In the second condition, a base word (frequency = 10 per million) and *five* derivations (frequency = 2 per million) were encoded. Notice that the token frequency across the derivations was equated so that any differences in the processing of the base forms can be attributed to the type frequency of their derivations. The base words' orthographic features were then used as probes to determine their familiarity and their pronunciation/meaning availability.

The simulation results in Panel C of Figure 2 show a clear effect of family size: The pronunciation and meaning of the base word with many derivations are more available than those of the base word with a single derivation. Interestingly, however, the type frequency of the derivational forms did not affect the overall familiarity of the base words—both base words were equally familiar even though derivational type frequency affected the availability of their identities. This result suggests an interesting and nonobvious disassociation between identifiability and familiarity that is not seen in monomorphemic words.

Summary. The morphology simulations suggest that our model can handle at least some phenomena in morphology. The simulations show a separation of inflectional and derivational effects based not on their predefined status but as a function of differences in their orthographic, phonological, and semantic similarity. Token frequency effects are restricted to morphological variations similar in form and meaning, a condition that applies more often to inflectional than derivational families. Type effects, however, are important for derivational families. Moreover, the token frequency affected both familiarity and availability of the base form. In contrast, type frequency affected the availability of the base word's identity but not its familiarity.

DISCUSSION

Our simulations demonstrate that basic memory processes and a few assumptions about the nature of word knowledge are sufficient to explain at least some important aspects of word identification, including the effects of frequency of occurrence and representational consistency. Our account of these variables connects two theoretical domains that—until now—have been largely disjoint: Those related to the identification of simple (monomorphemic) words and those related to sublexical (morphemic) processing in complex words. Our model suggests that these effects may share a common mechanism: the basic computational principles that allow knowledge about the spellings, pronunciations, and meanings of words to be encoded, represented, and retrieved from episodic memory. In what follows, we discuss some of the ramifications of these results for theories of morphology and models of word identification.

First, what can this approach add to the experimental research on morphology? Although this research is substantial, inconsistencies raise many questions. One problem is that many of the reported morpheme effects are quite subtle and fail to replicate with minor variations in procedures or materials. Our simulations suggest that variations in spelling, pronunciation, and meaning similarity are important and will produce variable outcomes within the same experimental manipulation. They also suggest that task differences are important because variation in word similarity can affect its familiarity (which may play a larger role in lexical decision than in other tasks) and the availability of its identity in different ways.

A second issue concerns the fact that theories of morphology have borrowed linguistic distinctions (e.g., derivation vs. inflection) but have not explained the cognitive mechanisms mediating these distinctions. In contrast, we have demonstrated how modeling can add to our understanding of how morphology is represented and used during reading. Our simulations address this question: How do the operating principles of a word-experience model of reading (based on episodic memory) influence the way morphemically complex words are represented and made available during reading? We do not assume that the distinction between inflectional and derivational forms is lost on language users when production is involved, or during tasks that tap people's awareness of meaning and form. Rather, we assume that a rapid process of skilled reading can take advantage of the form and meaning relations that are shared among words, irrespective of the kind of morphology that connects them.

A third issue concerns individual differences in reading ability. Our model instantiates a theory (Perfetti & Hart, 2001) of these differences in terms of *lexical quality*, or the degree to which the orthographic, phonological, and semantic features that collectively define a given word are both well represented and well interlocked in the reader's memory. This notion is captured by our model's assumption that word knowledge is acquired over time through repeated experiences with words. According to this view, individual differences in reading skill arise from differences in word knowledge, which in turn arise from differences in word experiences.[3] Thus, for skilled readers, word knowledge is well developed, with many words being represented in memory, and with many of these words being well represented. A high-quality representation is one that allows the reader to retrieve a word's spelling, pronunciation, and/or meaning from any one of these three sources of information. In addition, for many skilled readers, this information also includes some tacit knowledge about morphology—spellings, pronunciations, and meanings within both inflectional and derivational word families.

It is also important to note that word knowledge can be used in different ways. For example, whatever information about a word happens to be available (e.g., its spelling) can be used to evaluate its overall familiarity or to retrieve additional information (e.g., its pronunciation and meaning). Because this information is retrieved from long-term memory via a global-matching retrieval process, the exact information retrieved will depend on the frequency with which the information being sought was encoded and its similarity (and relative consistency) to the overall contents of long-term memory. By implementing a complete model (i.e., orthography, phonology, and semantics), we have demonstrated how these basic principles, in conjunction with our assumptions about how words are represented, can explain phenomena, like morphology, that have been ignored by existing models of word identification (Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Harm & Seidenberg, 1999; McClelland

[3]This may also reflect (and contribute to) motivational differences because skilled readers are more likely to enjoy (and hence participate in) reading than their less skilled peers (Stanovich, 2000).

& Rumelhart, 1981; Paap, Newsome, McDonald, & Schvaneveldt, 1982; Plaut, McClelland, Seidenberg, & Patterson, 1996).

Finally, although our model captures the results of prior experience on current performance, it may have some implications for issues of development and reading acquisition. The simple implication is that acquiring skill in reading requires encounters with words that build up representations that reflect familiarity and knowledge. This implication is in the spirit of models that emphasize the importance of reading practice for establishing the orthographic knowledge required for word-specific representations (Perfetti, 1992; Stanovich & West, 1989). The role of specific word experience is critical, of course, but important empirical questions remain: Beyond phonological decoding (e.g., Share & Stanovich, 1995), what learning experiences best promote the acquisition of form and meaning knowledge? Is some level of morphological awareness important to take maximal advantage of this experience? Our model emphasizes the importance of building up orthographic, phonological, and semantic information shared across words but also differentiates them. Experiences shape a reader's knowledge about language (e.g., phonology and morphology) and about printed word forms (orthography), and both contribute to the skilled model of reading that has been our focus.

ACKNOWLEDGMENTS

We thank Anna Bosman and Dominiek Sandra for their helpful suggestions on an earlier version of this article.

REFERENCES

Andrew, S. (1982). Phonological reading: Is the regularity effect consistent? *Memory & Cognition*, *10*, 565–575.

Balota, D., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340–357.

Balota, D., & Chumbley, J. I. (1985). The locus of the frequency effect in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, *24*, 89–106.

Berent, I., Pinker, S., & Shimron, J. (1999). Default nominal inflection in Hebrew: Evidence for mental variables. *Cognition*, *72*, 127–138.

Bertram, R., Baayen, R. H., & Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language*, *42*, 390–405.

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

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production, Vol. II: Development, writing and other language processes* (pp. 257–294). London: Academic.

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

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel distributed processing approaches. *Psychological Review*, *100*, 598–608.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.

Engbert, R., & Kliegl, R. (2001). Mathematical models of eye movements in reading: A possible role for autonomous saccades. *Biological Cybernetics*, *85*, 77–87.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamic model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, *42*, 621–636.

Francis, W. N., & Kuèera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Frederiksen, J. P., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 361–379.

Hare, M., Elman, J. L., & Daugherty, K. G. (1995). Default generalization in connectionist networks. *Language and Cognitive Processes*, *10*, 601–630.

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

Henderson, L., Wallis, J., & Knight, K. (1994). Morphemic structure and lexical access. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 211–226). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, *16*, 96–101.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.

Hintzman, D. L. (1987). Recognition and recall in MINERVA 2: Analysis of the "recognition-failure" paradigm. In P. Morris (Ed.), *Modeling cognition* (pp. 215–229). New York: Wiley.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528–551.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1612–1627.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40*, 431–439.

Jared, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory and Language*, *36*, 505–529.

Jarvella, R., & Meijers, G. (1983). Recognizing morphemes in spoken words: Some evidence for a stem-organized mental lexicon. In G. B. Flores d'Arcaos & R. Jarvella (Eds.), *The process of language understanding* (pp. 81–112). New York: Wiley.

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

Kliegl, R., & Engbert, R. (in press). SWIFT explorations. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eyes: Cognitive and applied aspects of oculomotor research*. Oxford, England: Elsevier.

Logan, G. D. (1988). Towards an instance theory of automatisation. *Psychological Review*, *95*, 492–527.

MacKay, D. G. (1978). Derivational rules and the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *17*, 61–71.

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

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

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

McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1155–1172.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.

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

Nosofsky, R. M. (1992). Exemplars, prototypes, and similarity rules. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning theory to connectionist theory: Essays in honor of William K. Estes* (pp. 149–167). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word-superiority effect. *Psychological Review*, *89*, 573–594.

Patterson, K., Seidenberg, M. S., & McClelland, J. L. (1989). Connections and dissociations: Acquired dyslexia in a computational model of reading processes. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neurobiology* (pp. 131–181). New York: Oxford University Press.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. M. Borwn & P. Hagoort (Eds.), *The neurocognition of language* (pp. 167–208). New York: Oxford University Press.

Perfetti, C. A., & Hart, L. A. (2001). The lexical bases of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection* (pp. 67–86). Washington, DC: American Psychological Association.

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

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 820–833.

Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, *8*, 1–56.

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review*, *97*, 285–308.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*, 385–408.

Rayner, J., & Duffy, S. A. (1986). Lexical ambiguity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*, 191–201.

Reichle, E. D., Landi, N., & Perfetti, C. A. (2002). *Toward an experience-based model of word identification*. Manuscript in preparation.

Reichle, E. D., & Perfetti, C. A. (2001). *Modeling lexical constituent availability during visual word identification*. Poster presented at the annual meeting of the Society for the Scientific Study of Reading, Boulder, CO.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). *Toward a model of eye movement control in reading*. *Psychological Review*, *105*, 125–157.

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E–Z Reader model. *Vision Research*, *39*, 4403–4411.

Reichle, E. D., Rayner, K., & Pollatsek, A. (in press). The E–Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*.

Rubin, G. S., Becker, C. A., & Freeman, R. H. (1979). Morphological structure and its effects on visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, *8*, 399–412.

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

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, *26*, 1270–1281.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language*, *37*, 118–139.

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, *19*, 1–30.

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

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, *23*, 383–404.

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of processing during reading: Evidence from eye movements and event-related potentials. *NeuroReport*, *9*, 2195–2200.

Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. *Issues in Education*, *1*, 1–57.

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford.

Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 658–672.

Stanovich, K. E., & West, R. F. (1983). On priming by sentence context. *Journal of Experimental Psychology: General*, *112*, 1–36.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24,* 402–433.

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes*, *9*, 271–294.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word recognition. *Journal of Memory and Language*, *26*, 608–631.

Tenpenny, P. L. (1995). Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, *2*, 339–363.

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time course and decision criteria. *Memory & Cognition*, *13*, 557–572.

Manuscript received June 6, 2002
Final revision received November 20, 2002
Accepted December 11, 2002

APPENDIX

MINVERVA 2 (Hintzman, 1984) consists of working memory (a vector of features representing the contents of consciousness) and memory traces (vectors representing the contents of memory). Features are perceptual and semantic primitives that take on values of +1, 0, and –1 (zero indicates absence) and include both *focal* information (e.g., a word's spelling and meaning) and *contextual* information (e.g., the physical setting in which the word is experienced).

Information is retrieved from memory via a global-matching process: Each feature *j* of a probe (i.e., the content of working memory) is compared in parallel to the features of each trace *i* to give the similarity between the probe and trace (using Equation 1, where N is the number of features and N_r is the total number of non-zero features in either the probe or trace).

$$(1) \quad \text{similarity}_i = [\sum_{j=1}^{N} (\text{probe}_j * \text{trace}_{i,j})] / N_r$$

Similarity values range from +1 (perfect match) to –1 (perfect mismatch). These values are then used to compute the activation generated by each trace *i*, using Equation 2. Traces very similar to the probe will generate disproportionately more activation (and contribute more to the information that is retrieved from memory) than those slightly less similar.

$$(2) \quad \text{activation}_i = \text{similarity}_i^3$$

The activation values are summed across all of the traces in memory to produce a signal, called the *echo intensity*, using Equation 3 (where M is the total number of traces in memory).

$$(3) \quad \text{intensity} = \sum_{i=1}^{M} \text{activation}_i$$

The echo intensity indicates how similar the probe is to all of the traces in memory, and is used to simulate recognition and frequency judgments. In our simulations, echo intensity is an index of a word's *familiarity*, indicating how well the word is established in a reader's vocabulary.

To simulate recall (which is necessary to access a word's spelling, sound, and/or meaning), a pattern-completion process is used to generate a pattern of activated features, or *echo content*. Each feature *j* of the echo content is found using Equation 4, which provides a basis of recall: If features of two "items" (e.g., a word's spelling and sound) are stored in the same trace, then probing with features of one item will produce an echo content that resembles the other.

(4) $\text{content}_j = \sum_{i=1}^{M} (\text{activation}_i * \text{trace}_{i,j})$

Because the echo content is a composite of many memory traces, it will often be degraded or noisy. In our simulations, the Pearson correlation coefficient, *r*, between the features of the echo content and those of the correct response was used as an index of the how well the echo content matched a word's spelling, pronunciation, or meaning (i.e., its *availability*).

Each trace of 2,400 features is divided into three equal-sized sets representing orthography, phonology, and semantics. The value of each feature *j* of word *n* was normalized (using Equation 5) to reflect its frequency of occurrence (Francis & Kučera, 1982) relative to other words. (This was done in lieu of encoding multiple tokens of each word.) One is added to the numerator of Equation 5 so that active features have nonzero values; one is added to the frequency of the most common word (*the*, which occurs 69,975 times per million words of printed text) so that all active feature values are in the –1 to +1 interval (excluding zero).

(5) $\text{feature}_{j,n} = \ln(\text{frequency}_n + 1) / \ln[\max(\text{frequency}) + 1]$

Finally, it was necessary to degrade the model's performance so that it would more closely resemble that of an adult reader with a vocabulary of tens of thousands of words. This was done by adding a "noisy" memory trace having features that were random deviates independently sampled from a Gaussian distribution with $\mu = 0$ and $\sigma = 5$. As expected, this precaution markedly reduced the model's overall performance.