Connectionism and the Role of Morphology in Visual Word Recognition

Jay G. Rueckl
University of Connecticut and Haskins Laboratories

Abstract

This paper provides a review of the connectionist perspective on the role of morphology in visual word recognition. Several computational models of morphological effects in reading are described and relationships between these models, models of past tense production, and models of other aspects of word recognition are traced. Limitations of extant models are noted, as are some of the technical challenges that must be solved to develop the next generation of models. Finally, some directions for future research are identified.

Readers are influenced by the morphological structure of the words that they read. Theories attempting to account for this fact have generally supposed that the impact of morphology on reading comes about because readers explicitly represent the morphological structure of written words. Although the details of such theories vary (cf. Diependaele, Sandra, & Grainger, 2005; Frost, Forster, & Deutsch, 1997; Rastle & Davis, 2008; Taft, 1979, 1994), what they have in common is the assumption that at least some morphologically complex words are decomposed into their constituent morphemes, and thus that there is an explicitly morphological level of representation.

The connectionist framework provides a different perspective—one that views morphological effects not as the result of the structural properties of the lexicon, but instead as the influence of statistical regularities in the mappings between (orthographic and phonological) form and meaning. On this view, information about formal and semantic similarity converges on a common set of processing units. As a consequence, over the course of learning the patterns of activation over these units come to capture 'morphological' structure.

One purpose of this article is to explicate the connectionist perspective on morphological effects in word recognition and review several computational modeling efforts embodying it. In part, this entails reviewing related work on past tense production and other aspects of visual word recognition. Another purpose of this article is more forward-looking—after reviewing extant models I discuss some of the technical issues that must be solved to develop the next generation of models, as well as some directions for future research suggested by these and other considerations

Connectionism: Theory and Applications

Connectionist models are sometimes called 'neurally inspired' because they incorporate structures and processes that are meant to mirror—at a quite abstract level —those found in the brain. Thus, a connectionist network is composed of many simple, neuron-like

Address correspondence to: Jay Rueckl, Department of Psychology, Box 1020, University of Connecticut, Storrs, CT 06269, Fax: (860) 486-2760. Jay.Rueckl@uconn.edu.

processing units called *nodes* that communicate by sending excitatory and inhibitory signals to one another. Each signal is weighted by the strength of the connection that it is sent across, and the state of each node (its *activation*) is a nonlinear function of the sum of these weighted signals. Like neural synapses, the connections in a network are plastic, and a learning algorithm is used to adjust their strengths (or *weights*) such that, over the course of learning, the flow of activation becomes tailored to the structure and task demands of the environment in which the network is embedded. (For overviews, see Elman et al., 1996, or Rumelhart et al., 1986.)

While this general framework is essential to the connectionist approach, a connectionist model is more than a specification of the activation and learning algorithms. First, in connectionist models, cognitive tasks are described as mappings from inputs to outputs. Thus, building a model entails a specification of the input and output domains. Are past-tense forms generated on the basis of phonological or semantic inputs? Are written words mapped onto semantic representations or abstract (and meaningless) lexical representations? Second, a network learns to perform a mapping through its interactions with its task environment. Thus, the structure of this mapping must be specified. What are the input-output correspondences that the network will be exposed to? Are some inputs more frequent than others? Does the distribution of inputs change over the course of learning? Third, the interface between the network and its environment—namely, the representations of its inputs and outputs—must be specified. These representational schemes determine the similarity metrics within the input and output domains, and hence whether statistical regularities are or are not available in the network's training environment. Finally, the architecture of the network must also be specified. Does the network map its inputs directly onto its outputs, or is this mapping accomplished via an intermediate (hidden) layer of nodes —allowing the network to take advantage of learned internal representations?

Before turning to models of morphological effects in word recognition, it will be instructive to consider how these issues play out in the development of two families of connectionist models—models of past-tense production and models of visual word recognition. In addition to exemplifying the connectionist approach, the evolution of these models provides some guideposts for the next generation of models.

Past tense

The publication of Rumelhart and McClelland's (1986) model of past-tense production (hereafter, the RM model) was a landmark event in computational modeling and in cognitive science more generally. In some respects, the production of English past-tense forms is an unlikely candidate for the theoretical attention it has received. The acquisition of the past tense is a relatively minor aspect of language acquisition as a whole. Moreover, compared to many other languages (e.g., Finnish, Hebrew, Russian), English is rather impoverished morphologically, and thus affords relatively limited opportunities for experimentation.

So why did the past tense model garner the attention it did? For one thing, the manner in which children acquire the English past tense has been taken as strong evidence for a pervasive 'rules plus memory' approach to language processing. As described by Brown (1973) and others, the acquisition of the English past tense can be characterized as a sequence of three stages: An initial stage, during which children produce the past tense of a small number of regular and irregular verbs, an intermediate stage, during which the number of past-tense forms that are generated grows rapidly, children begin to exhibit the ability to generate past-tense forms for novel verbs (e.g., *rick-ricked*, Berko, 1958) and also sometimes *overregularize* irregular forms (i.e., produce '*comed*' or *camed*' rather than '*came*'; Ervin, 1964; Kuczaj, 1977), and a final stage in which difficulties with irregular forms diminish and the child becomes proficient in producing both regular and irregular

forms (although even adults overregularize on occasion and sometimes exhibit uncertainty about the correct form of some low-frequency verbs). According to the canonical "rule + memory" account, during the first stage a small number of forms are simply memorized. Behavior during the second stage reveals the discovery of the 'add –ed' rule. This rule not only allows for a rapid expansion in the number of familiar past-tense forms that can be produced, but can also be used to generate past-tense forms for novel items. However, it also generates incorrect forms for irregular words. Thus, the overregularization errors observed during Stage 2 are thought to result from the misapplication of the rule. With additional learning irregular forms become better established in memory, and hence in Stage 3 overregularizations occur relatively infrequently.

Thus, the acquisition of English past tense has been taken as both a window on the learning of a linguistic rule and a demonstration of how two fundamentally distinct computational mechanisms come to jointly determine a complex pattern of behavior. The RM model challenged both of these assumptions. In the RM model, rule-like behavior emerges from the dynamics of a computational mechanism that neither represents nor applies a rule; moreover, the same mechanism is responsible for the production of both regular and irregular forms. The RM model characterizes the acquisition of the past tense as the problem of learning to generate a representation of the phonological form of a past-tense verb given as an input a representation of the phonological form of the corresponding verb stem. The model includes two layers of neuron-like processing units.[i] Patterns of activation over the input layer represent verb stems, with the phonological similarity of the stems reflected by the similarity of the corresponding patterns of activation. Patterns of activation over the output layer represent past-tense forms. Here too, phonologically similar forms are represented by similar patterns of activation.

Rumelhart & McClelland trained their network with a training corpus that was meant to be representative of a child's environment. Thus, early in training the network was exposed to a small number of high-frequency regular and exception words. As performance on these words improved, the training set was expanded to include a larger set of words, most of which take regular inflections. Because the weights were initially set to random values, the network could not generate any correct responses before training began. By the end of training, however, the network's performance on both regular and irregular forms was highly accurate, thus demonstrating the feasibility of the single-mechanism approach. Critically, while performance on regular verbs increased steadily over the course of learning, the production of irregular forms exhibited the characteristic 'U-shaped' function: At an intermediate point in training the network began to make mistakes on irregular forms it had previously been able to generate; with further training the network began to again produce the correct forms. Moreover, when the network made a mistake, the output was often the sort of overgeneralization that would result from the misapplication of the add –ed rule. In addition, when tested with unfamiliar words at the end of training, the network produced plausible responses to most of the test items. (The responses typically conformed to the add –ed rule, although in a few cases novel items that were similar to a cluster of irregular words resulted in an irregular past-tense form.)

Based on these results, Rumelhart and McClelland wrote that "We have, we believe, provided a distinctive alternative to the view that children learn the rules of English past-tense formation in any explicit sense. We have shown that a reasonable account of the acquisition of past tense can be provided without recourse to the notion of a "rule" as anything more than a *description* of the language" (1986, p. 267). Of course, not everyone

[i]Technically, the model included four layers of nodes. However, because only the connection weights between two of these layers were allowed to change with learning, it is more appropriate to think of the model as a two-layer network.

agreed. The RM model met with stiff opposition (e.g., Lachter & Bever, 1988; Marcus, Brinkman, Clahsen, Wiese, & Pinker, 1995; Marslen-Wilson & Tyler, 1998; Pinker & Prince, 1988,), which in turn sparked both further development of the model and a variety of new empirical studies. These theoretical and empirical developments are extensive and largely beyond the scope of this article (see McClelland & Patterson, 2003, Pinker & Ullman, 2003, and Seidenberg & Gonnerman, 2000, for comprehensive reviews). However, several issues are especially relevant to models of morphological effects in reading, and hence will be highlighted here. These include questions about the level of performance of the RM model and the degree to which its behavior mirrored that of children acquiring English (including the fact that on rare occasion it produced unusual outputs like 'membled'), the tight coupling between the 'vocabulary discontinuity' in its training environment and the onset of overgeneralization errors, and its inability to deal with homophony (e.g., *ring/rang, wring/wrung*).

The original RM model included a single set of adjustable connections. The computational limitations of this kind of pattern associator are well known (Minsky & Papert, 1969) and can be overcome by using the sort of learning algorithms that began to emerge around the time that the RM model was published. Most prominent among these is the backpropagation algorithm (e.g. Rumelhart, Hinton, & Williams, 1986), which provides a means of adjusting the strengths of connections to and from a layer of 'hidden' units that mediate the mapping between the input and output representations. The representations over the hidden layer are not stipulated by the modeler, but instead are organized by the learning process in response to the demands of the task and the structure of the training environment. Plunkett and Marchman (1991) contrasted trained networks with and without hidden units to produce English past tense forms. Their findings (substantiated in a number of subsequent simulation studies) demonstrated that the inclusion of hidden units improved the overall performance of the network and resulted in patterns of behavior more in keeping with what has been observed in empirical studies.

Another response to the criticisms of the RM model was to explore the relationship between the structure of the training environment and the network's behavior. These simulations demonstrated that an external discontinuity in the structure of the training environment was neither a necessary nor sufficient condition for the occurrence of U-shaped learning curves. U-shaped curves have been found when the size of the training set increases gradually over the course of training (Plunkett & Marchman, 1991, 1993) and with a training set that is fixed over the training regime (Cottrell & Plunkett, 1994). But this is not to say that the structure of the training environment is irrelevant. Various computational modeling studies have demonstrated that the behavior of a network trained to produce past tense forms is influenced by the size of the training set, the distribution of different (regular and irregular) verb types, and the relative frequency of individual verbs (Plunkett & Marchman, 1991, 1993; Hahn & Nakisa, 2000). Moreover, a network's behavior is also contingent on the appropriateness of the input and output representations specified by the modeler—given the wrong choices, the input-output mapping performed by the network may not capture the statistical regularities ascribed to the language (MacWhinney and Leinbach, 1991).

A third key development was the reformulation of past tense production as a task that maps semantic representations to phonological representations (Cottrell & Plunkett, 1994). One advantage of this formulation is that it addresses the question of homophony: *ring* and *wring* are distinguishable (and hence can be mapped onto *rang* and *wrung*) because they have different meanings. On the other hand, if the network's input is semantic, it has no way to generate the past tense of nonce forms such as *wug*. This shortcoming was remedied by Joannise and Seidenberg (1999), who trained a network to generate phonological output representations (of both present and past tense forms) from phonological input

representations as well as semantic representations. Both mappings were mediated by a common set of hidden units, and hence the network was not simply the juxtaposition of two independent processes. Various analyses revealed that although both regular and irregular past tense forms could be generated from either type of input, the semantic and phonological components of the model played somewhat different roles. Thus, while damage to the phonological input representations impaired the production of nonce forms more than the production of irregular forms, the reverse was true when the semantic system was damaged.

The above extensions of the RM model are important not only because they sharpened the connectionist account of past tense production, but also because they illustrate several of the core tenets of the connectionist approach. One of these is the importance of statistical structure. Learning attunes a network to the structure of its environment, and the goal a computational model is to understand how patterns of behavior emerge from the interaction of a network and its environment. Thus, the value of a model depends in part on the degree to which its learning environment does (or does not) capture the statistical regularities that are relevant to human behavior. A second tenet is that behavior reflects the organization of the learned internal (hidden) representations that determine a network's response to an input. Computational modeling can be used to reveal how this organization comes about and how it determines the organization of behavior. Finally, the Joanisse and Seidenberg (1999) model illustrates a third tenet. In general, networks are not homogeneous wholes, but instead are composed of constituent subsystems that are differentially sensitive to different kinds of task constraints. The behavior of a network is the product of cooperative interactions among these component subsystems. Thus, there is specialization, but not modularity. Computational modeling can be used to elucidate the division of labor among a network's components as it performs a cognitive task.

Visual Word Recognition

Shortly after the publication of the RM model, Seidenberg and McClelland (1989; hereafter SM89) reported the first implementation of what would come to be called the 'triangle model'. Like the RM model, the triangle model offered a new view of how language users deal with quasi-regularity—in this case, quasi-regularity in the mapping between written and spoken word forms. In alphabetic writing systems, the mapping from orthography to phonology can largely be captured by a set of rules that specify how to convert a string of letters into a sequence of phonemes. However, in many writing systems there are exceptions to these rules. For example, in English the 'grapheme-phoneme conversion rules' that specify the pronunciation of words like MINT and SAVE generate incorrect pronunciations for other words, such as PINT and HAVE.

Like the traditional account of how language users deal with irregularities in the past-tense system, models of visual word recognition have typically adopted a 'rules plus memory' solution to deal with this kind of quasi-regularity. According to such accounts, readers employ two distinct processes: a process that retrieves lexical information directly from a word's spelling (the 'memory' route) and a sublexical process involving the computation of an phonological code (the 'rule' route) which can in turn be used to access word meaning. In the standard interpretation, *exception* words (PINT, HAVE) must be read via the lexical route, nonwords (CRINT, MAVE) can only be read via the sublexical route, and *regular* words (MINT, SAVE) can be read by either route. An important empirical observation in support of this view is that regular words are read faster than exception words (Barron, 1976; Glushko, 1979). According to so-called "dual-route" models, the regularity effect is a consequence of the fact that only regular words benefit from the availability of the phonological route. As a word becomes more familiar, the phonological route contributes less to its processing—hence the observation that the regularity effect is largest for low-

frequency words (Andrews, 1982; Seidenberg, 1985;Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Taraban &McClelland, 1987; Waters & Seidenberg, 1985).

Like other models of word recognition, the triangle model embraces the idea that the organization of the reading system reflects the fact that writing systems afford two ways to compute word meaning (directly from spelling of via an intermediate phonological code). Thus, the model includes distinct layers of nodes responsible for representing the orthographic, semantic, and phonological properties of written words, with separate sets of connections (and hidden units) mediating each of the three (O-S, O-P and P-S) mappings. In each layer, the patterns of activation that represent different words are organized such that words that are similar on the relevant linguistic dimension are represented by similar patterns of activation. As in the RM model (and its descendents), an incremental learning algorithm attunes the network to the structure of the mappings among these representations.

The primary achievement of the SM89 implementation of the triangle model was the demonstration that the ability of skilled readers to cope with quasi-regular nature of the mapping from orthography to phonology does not entail the existence of two distinct mechanisms, one to capture the regularities and one to handle the exceptions. The SM89 model focused on the O-P leg of the triangle, with the O-S and P-S pathways left unimplemented. The model was trained on a moderately sized (~3000 word) corpus that captured key aspects of the structure of English spelling-sound correspondences (e.g., quasi-regularity, homophony, variation in word frequency, and so on). Although Seidenberg and McClelland provided a rich analysis of the network's behavior and internal organization, only two key results will be highlighted here. First, despite the fact that it used the same mechanism to read regular and exception words, the network exhibited the classic regularity x frequency interaction that has been interpreted as revealing the operation of two distinct mechanisms. Second, the same mechanism underlying the naming of familiar words also proved capable of naming unfamiliar nonwords, again demonstrating how a network can generalize the knowledge it acquires as it becomes attuned to the structure of its task environment.

Like the original version of the past tense model, the initial SM89 implementation of the triangle model met with strong opposition (e.g. Besner, Twilley, McCann, R., & Seergobin, 1990; Coltheart, Curtis, Atkins, & Hailer, 1993), and as in the case of the past tense model, the resulting debate led to both refinements of the model (Harm & Seidenberg, 1999; Harm et al., 2003; Kello & Plaut, 2003; Plaut, 1995; Plaut & Shallice, 1993) and new lines of empirical inquiry (e.g., Seidenberg et al., 1996; Spieler & Balota, 1997; Treiman et al., 2003, Woollams, Lambon Ralph, Plaut, & Patterson, 2007). Again, much of this work is beyond the scope of this paper (see Coltheart et al., 2001, Harm & Seidenberg, 2004, Seidenberg & Plaut, 2006, for reviews), but several particular developments are worth highlighting.

First, one of the criticisms of the SM89 model is that its ability to read nonwords was not on a par with its performance with familiar words (Coltheart et al., 1993). In a follow-up to the Seidenberg and McClelland simulations, Plaut et al. (1996) noted that the orthographic and phonological representations used in SM89 created a 'dispersion problem' —spelling-sound correspondences were distributed across a number of input and output units, making it difficult for the network to become adequately attuned to these regularities. Plaut et al. demonstrated that orthographic and phonological representations that concentrate these correspondences on a smaller number of units improve the network's ability to read nonwords without impairing its performance with familiar words. This improvement in the performance again illustrates the importance of statistical structure—the behavior of the

model is determined by the regularities embodied in its training set. Get the regularities wrong and the model's behavior will reflect it.

A second development that emerged in the follow-ups to the SM89 model was a deeper understanding of the organization of the hidden representations that mediate the O-P mapping. One important insight is that the hidden representations are as *componential* as the prevailing conditions allow. That is, if there are statistical regularities involving components of an input (e.g., letters, word bodies), the patterns of activation over the hidden units representing that input will contain (more or less) subpatterns corresponding to these components (Plaut et al., 1996). Another key insight is that the hidden representations are organized to capture both similarities among the input patterns and similarities among the responses to which these inputs must be mapped (Harm et al., 2003). Thus, for example, the hidden units mediating the O-P mapping are organized such that words that are similar in spelling (LAKE, TAKE) have relatively similar hidden representations, but so too do words that are similar in pronunciation (BEAR, BARE). Taken together, these characteristics provide the network with an efficient means for dealing with quasi-regularity. For example, by positioning the representation of PINT somewhere near—but not too near—the representations of MINT and HINT, the network can take advantage of the similarities among these words (namely, that NT is pronounced/nt/) while also ensuring that PINT isn't pronounced as a rhyme of MINT. Similarly, learned hidden representations also provide a means for the network to generalize its knowledge to novel situations—e.g., generating a plausible pronunciation of the nonword ZINT.

A third advance in our understanding of network models concerns the division of labor among a network's components. A major development in this regard was the set of simulations reported by Harm and Seidenberg (2004). Unlike most of the previous instantiations of the triangle model, the Harm and Seidenberg version included all three sides of the triangle (i.e., the O-P, O-S, and P-S components). Using a variety of measures, they demonstrated that semantic activation depends on the cooperative interactions of the O-S and O-P-S pathways, and conversely, that the activation of a phonological code depends on both the O-P and O-S-P pathways. These measures also revealed that although reading involved the cooperative interactions of the O-S and O-P components, the division of labor between these components varied as a function of various lexical properties (e.g., frequency, spelling–sound consistency, homophony) and changed over the course of development. (Early reading was relied largely on the O-P component, but as learning progressed and the O-S pathway became more efficient, the model moved towards a more cooperative division of labor.)

In sum, advances in both the past tense model and the triangle model have highlighted the importance of statistical structure and provided insights about the organization of a network's hidden representations and the division of labor between its component subsystems. These lessons set the stage for a discussion of network models of morphological effects in visual word recognition.

Morphological Structure and Visual Word Recognition

From a connectionist perspective, the effects of morphological structure on word recognition arise because statistical regularities related to morphology influence the dynamics of the processes that map representations of orthographic and phonological form to representations of meaning (Rueckl et al., 1997; Seidenberg & Gonnerman, 2000). Given the prominent role of this position in motivating experimental investigations of morphological effects in reading, it is perhaps surprising that the number of computational modeling studies addressing such phenomena is actually rather small.

The first of these simulation studies was reported by Rueckl and Raveh (1999). The motivation of this study was the observation that morphological structure creates islands of regularities in the otherwise arbitrary mapping from form to meaning. That is, in contrast to the mappings from stem to past tense and spelling to pronunciation, where similar inputs are usually mapped to similar outputs, in the mapping from form to meaning similar forms are typically not related in meaning. (Consider *make, take, lake*, and *wake.*) Morphologically related words are an exception to this arbitrariness.

Because the most prominent connectionist models at the time (namely, models of past-tense acquisition and models of word naming) had investigated the influence of statistical regularities in domains with highly systematic mappings, the question of whether a network can exploit regularities that occur against the backdrop of an otherwise unstructured mapping, or whether instead the potential influence of such regularities is overpowered by the arbitrary character of the mapping as a whole, remained unresolved. To address this question, Rueckl and Raveh compared the behavior of identical networks trained on two kinds of mappings.

For some networks, the mapping from orthographic input patterns to semantic output patterns was structured by morphological regularities. To create this mapping, morphological families were generated by creating 3-letter stems and concatenating each stem with each of three 1-letter suffixes. Each stem was assigned a 'meaning' by pairing it with a randomly determined set of semantic features. The meaning of an inflected form was represented by the activation of the semantic pattern associated with the stem, together with the activation of an additional output node that was uniquely and consistently paired with a given suffix. Thus, although the mapping from spelling to meaning was largely arbitrary (due to the process used to assign meanings to the stems), morphological relationships imparted some structure to this mapping. For other networks the words and meanings were randomly re-paired, ensuring that the input-output mapping was completely arbitrary.

The results revealed that networks trained on mappings structured by morphological regularities learned more quickly and were capable of learning larger vocabularies than networks trained on completely arbitrary mappings. Various analyses demonstrated that these behavioral effects arose because the internal representations of networks trained on morphologically structured mappings were shaped by these regularities: Similarly spelled words were represented by similar hidden patterns, and this was especially true for morphologically related words. Moreover, these patterns had a componential structure such that the pattern for an affixed word corresponded (approximately) to the superimposition of subpatterns corresponding to each of its morphological constituents.

The Rueckl and Raveh (1999) simulations were intended to illuminate a general property of connectionist networks but were not meant to capture and specific empirical phenomena. In contrast, in a study that was published the following year, Plaut and Gonnerman (2000) focused on experimental results demonstrating that morphological priming varies with semantic transparency. For example, in a cross-model priming study, Gonnerman et al. (2007) found more priming for highly related prime-target pairs (BAKER–BAKE) than for moderately related pairs (DRESSER–DRESS), which in turn produced more priming than semantically unrelated pairs (CORNER–CORN).

To model these kinds of priming effects, Plaut and Gonnerman (2000) trained a network on an artificial language constructed using a strategy similar to that used by Rueckl and Raveh (1999)—albeit differing in detail. Plaut and Gonnerman's training set included larger morphological families and, most critically, the semantic relationships among the words in a family were manipulated by distorting (to differing extents) the canonical semantic patterns

associated with each stem and suffix. In addition, to model cross-language differences in morphological 'richness' (e.g., the contrast between 'rich' languages like Finnish or Hebrew and 'impoverished' languages like English), the global structure of the form-meaning mapping was manipulated by varying the proportion of semantically unrelated words containing the same orthographic stem. Plaut and Gonnerman trained identical networks on these mappings. At the end of training, an approximation of a continuous-time activation rule was used to study the effects of semantic transparency on morphological priming. On a given trial, the network was given an input (the prime) and activation was allowed to flow through the network for a fixed amount of time (such that, in effect, the prime had been partially processed). Then the input corresponding to the target was presented (without altering the activations of the other units in the network) and activation continued to flow until the network had settled into a stable state. The time needed to reach a stable state is taken as the analog of response time in a priming experiment, and the question of interest is whether and how response times vary as a function of the relationship between the prime and target.

Three results are of particular interest. First, relative to an (orthographically and semantically) unrelated control condition, the network settled more quickly when the primes and targets were morphologically related. This effect occurred because (as was also found by Rueckl & Raveh, 1999) the hidden-unit representations were organized to capture morphological similarity. Thus, morphologically related primes and targets are represented by similar hidden patterns, and seeing a related prime not only moves the system towards the pattern representing the prime, it also moves the system towards the pattern associated with the target.

The second key finding is that semantic transparency modulated this priming effect: Facilitation decreased monotonically as a function of the semantic distance between the prime and target. Interestingly, in this case the magnitude of priming was only loosely related to the similarity of the hidden patterns. In particular, although the hidden patterns for pairs of semantically unrelated words were less similar than those for pairs of semantically related words (with orthographic overlap held constant), there was no systematic effect of *degree* of relatedness among the semantically related pairs. This suggests that the semantic transparency effect reflects the degree to which the prime directly activates the semantic features of the target (i.e., the degree to which it activates the target's *output* representation, rather than its hidden representation), although another possibility is the metric used to assess hidden-pattern similarity (in this case, Pearson correlations) fails to pick up an important dimension of variability.

The third key finding is that varied with the global structure of the mapping from spelling to meaning. Priming effects were generally larger for the network trained on the morphologically rich language, and although priming was graded as a function of semantic transparency in both languages, the drop-off was more pronounced in the impoverished language. Plaut and Gonnerman noted that this pattern is similar to what has been observed experimentally (cf. Bentin & Feldman, 1990; Frost et al., 2000; Feldman & Soltano, 1999; Marslen-Wilson et al., 1994) and concluded that the degree to which morphologically related words are represented by similar patterns depends in part on the global structure of the form-meaning mapping.

The Rueckl and Raveh (1999) and Plaut and Gonnerman (2000) simulations involved training sets that captured theoretically critical aspects of the mapping from spelling to meaning but were not based on any particular language. While this 'artificial language' approach allows the modeler to sidestep certain technical and implementational challenges (see below), it is fair to wonder whether the results would hold if the network were

confronted with a training set containing more of the regularities and idiosyncrasies inherent in a natural language. In this respect, simulations involving training sets grounded more in an extant language are particularly appealing.

One such set of simulations was reported by Harm and Seidenberg (2004), who implemented the full triangle model and trained it on a corpus of nearly 6000 words (virtually all the monosyllabic words in English). The primary goal of these simulations was not to model morphological effects in word recognition, but rather to examine the division of labor between processes that map printed words to their meanings directly or via an intermediate phonological code (i.e., the O-S and O-P-S pathways). However, the training set happened to include a number of morphologically complex forms—primarily plurals (DOGS, CATS, GEESE), past-tense forms (BAKED, SPARED, CAME), and third-person singular inflections (BAKES, SPARES, COMES)—in sufficient numbers to allow the network to become attuned to this (limited) set of morphological regularities.

Harm and Seidenberg investigated the network's sensitivity to morphological regularities by presenting it with nonwords such as GOME, GOMES, and GOMED. Inspection of the patterns of activation over the semantic layer revealed that 'inflected' nonwords (GOMES, GOMED) resulted in the strong activation of the appropriate semantic feature(s) (*plural* and *third person* in the case of GOMES, *past tense* in the case of GOMED), whereas the activation of other semantic features was far lower. Moreover, whereas the *past tense* feature was always nonwords ending in –ED, the *plural* and *third person* features were activated less consistently by nonwords ending in –S. This pattern is in essence a consistency effect, reflecting the greater variability in meaning of word-final –S.

To shed light on how the network makes use of morphological regularities, Harm and Seidenberg compared the behavior of the intact network with the operation of the isolated O-S and O-P-S pathways. (This was accomplished by 'lesioning' one or the other pathway after the network had been trained.) The performance of the isolated O-S pathway was nearly (but not exactly) identical to that of the intact network, indicating that the network had in fact become attuned to sublexical regularities in the O-S mapping. The isolated O-P-S pathway was less effective in activating the semantic features associated with each suffix, most likely as a consequence of variability in the phonemic realization of the printed suffixes (e.g.,—S is pronounced differently in LAKES and HANDS). This being noted, it would be a mistake to conclude that the influence of morphological structure in the model arises solely through the operation of the O-S pathway. The broad implication from the entirety of Harm and Seidenberg's simulations is that the reading involves the cooperative interaction of the O-S and O-P-S pathways. Thus, even if the structure of the O-S and O-P-S mappings is such that morphological regularities are more reliable in the O-S mapping, if the O-P-S pathway is available (as it is in the intact system) its operation will also contribute to the sensitivity to morphological structure manifest in the network's behavior.

Moving Forward: Technical Issues and Problems to be Solved

From a connectionist perspective, the effects of morphological structure on word recognition arise because statistical regularities related to morphology influence the dynamics of the processes that map representations of orthographic and phonological form to representations of meaning. Given the strong links between this idea and previous modeling efforts concerning both visual word recognition and past-tense production, and given too the prominent role of this view in motivating experimental investigations of morphological effects in reading, it is perhaps surprising that the number of computational modeling studies addressing such phenomena is actually rather small. On the other hand, given some of the technical challenges that must be addressed—particularly if the training corpus is to be

based on a natural language—perhaps this situation isn't surprising at all. In this section five such challenges are identified.

Semantics

As noted above, to implement a model the modeler must specify how a network's inputs and outputs are to be represented. Thus, to model morphological effects in reading, the modeler must decide how to represent both how a word is spelled and what it means. Developing a reasonably plausible scheme for representing orthography is not difficult (although see below); in contrast, the representation of word meaning is much more challenging. Plaut and Gonnerman (2000) and Rueckl and Raveh (1999) finessed this problem by using artificial languages, and thus simply generated 'semantic' patterns according to a convention that imposed the desired structure on the O-S mapping. Joanisse and Seidenberg (1999) followed a similar strategy in their past-tense model. In contrast, in their full-blown implementation of the triangle model, Harm and Seidenberg (2004) based their training set on monosyllabic English words, and thus needed to devise a scheme that captured the semantic relationships among these words in a reasonably accurate way. To generate semantic representations in a (quasi)algorithmic way, they developed a system (described in Harm, 2002) to extract the semantic features from WordNet, a computational thesaurus that provides semantic classification of the English lexicon in terms of hyponyms, synonyms, and antonyms. This process yielded 1,989 semantic features that were used to encode the meanings of over 6,000 words. (The number of features used to encode a word ranging from 1 to 37; hence the representations were rather sparse).

WordNet is one of several online databases that can be used for this purpose. Two others are Latent Semantic Analysis (LSA; Landauer & Dumais, 1997) and the Hyperspace Analog to Language (HAL; Burgess & Lund, 1997), both of which use computational algorithms to extract meanings from co-occurrence statistics from large-scale language corpora. There are several important advantages to deriving semantic representations from databases of this sort: The process is fairly automatic and can (in principle) be applied to any language. On the other hand, the translation from database to training set requires both decisions and technical skill on the part of the modeler. Moreover, the resulting feature sets may require some hand tuning. (For example, Harm & Seidenberg added features for *past tense*, *plural*, and *third-person singular*.) Third, it is unclear whether co-occurrence statistics capture the semantic relationship among morphologically related words (which are often in different syntactic categories) as we well as semantic relationships among morphologically unrelated words.

Scale

The quality of a model depends on the degree to which the structure of its training set mirrors the structure of person's environment. Because the relevant regularities are at a relatively small grain size, a corpus of 3000 to 6000 words appears sufficient to capture the distribution of spelling-sound correspondences (at least for monosyllable words). In contrast, because the number of morphemes in a language is far larger than the number of letters or phonemes, and too because the number of morphemes per word is less than the number of letters or phonemes per word, a much larger training set may be needed to capture morphological regularities in the mapping from form to meaning. Coupled with the large number of features needed to represent word meaning and the difficulty inherent in learning a relatively arbitrary mapping, the need for a large training set implies a computational scale far beyond that of any extent connectionist model. (For a step towards a large-scale model, see Sibley, Kello, & Seidenberg. 2010).

Masked Priming

Much of the recent research on morphological effects in visual word identification has involved the masked priming paradigm (Forster, Davis, Schoknecht, & Carter, 1987). In this paradigm a prime is presented for a very brief duration (typically 30–60 ms) and is both preceded and succeeded by masking stimuli (the pre-mask might be a row of #'s; the post-mask is often the target stimulus). Given the short duration of the prime, the fact that participants are often unable to report whether the prime was even present, and the pattern of results that are typically found, the masked priming paradigm has become a popular tool for studying the 'front end' of the reading system—the processes involved in the initial coding of the linguistic properties of a printed word. These early processes are thought to be primarily concerned with processing the identities and positions of the constituent letters of a word, although masked priming results suggest that these processes are sensitive to the morphological structure of a word as well (Feldman et al., 2009; Frost et al., 1997; Rastle et al., 2004).

From a modeling perspective, the findings from masked-priming studies pose two challenges. First, priming effects in this paradigm are quite sensitive to subtle methodological variations that seem to have much more to do with general visual processes than with word recognition per se (cf. Frost, Ahissar, Gotesman, & Tayeb, 2003; Michaels & Turvey, 1979). This raises the worry that the outcome of a simulation could be determined by assumptions that are secondary to the theory but are needed to model the priming paradigm. The slippery slope here is that to get the model's behavior to align with experimental results one would end up 'modeling the task' rather than modeling the word recognition process—an unattractive outcome for many of us who do modeling.

The second and theoretically more important challenge stems from the conclusion that the front end of the reading system is sensitive to morphological structure. In the triangle model the front end of the reading system is the input layer. The representations in this layer are stipulated by the modeler. Several different schemes have been used in the various implementations of the model (c.f., Seidenberg & McClelland, 1989; Plaut et al., 1996; Harm & Seidenberg, 2004), each of which was designed to capture information about letter identity and letter position (but not morphology). Of course, other schemes could be employed instead. For example, the input representations could be designed to explicitly represent morphological structure. This would have the advantage of putting morphology in the front end of the model, but at the cost that morphological structure would be stipulated by the modeler rather than discovered by the model.

A better approach would be to bring the implementation of the model into closer alignment with the theory. Past implementations of the model have employed stipulated orthographic representations in order to investigate the properties of the 'downstream' processes that map printed words onto semantic and phonological representations. However, the theory holds that all of the representations in the reading system are learned. Thus, what is really needed is a model in which the input units represent visual (rather than orthographic) properties of the input. The task for the network would be to learn to re-represent this input pattern in a manner that is appropriate for mapping it to the appropriate semantic and phonological codes. The expectation is that the resulting (learned) representations would capture information about letter identity and letter position, but would also be shaped by morphological (and phonological) regularities.

Future Directions

The previous section focused on some of the technical issues that need to be confronted to move the modeling of morphological effects in word recognition forward. The next section

considers the directions that future research should take given these considerations. These directions involve both implementations of the model and experimental research.

Morphological Decomposition

One of the major questions driving research on morphological effects in reading concerns when and how morphologically complex words are parsed into their constituent morphemes. A large body of evidence suggests that morphological effects arise early in the time course of word recognition. For example, as noted above morphological facilitation in the masked priming paradigm (e.g., Feldman & Soltano, 1999; Rastle, Davis, & New, 2004) is taken to indicate that morphology influences the front end of the reading process. The same conclusion can be drawn from the modality-specific component of long-term morphological priming (Rueckl & Galantucci, 2005; Rueckl & Aicher, 2008) as well as findings based on functional magnetic resonance imaging (fMRI, Devlin et al., 2006; Gold & Rastle, 2007) and electrophysiological measures (Morris et al., 2007).

These results are widely taken as evidence that morphological decomposition occurs pre-lexically, with representations of a word's constituent morphemes serving as the 'access units' to the mental lexicon (but see Giraudo & Grainger, 2001, and Seidenberg & Gonnerman, 2000, for alternative perspectives). Several hypotheses about the nature of the process that extracts morphological structure from the visual input have been proposed. For example, Rastle and Davis (2008; also see Rastle et al., 2004) have argued that decomposition occurs via a 'semantically blind' process that parses a word into morphemic constituents whenever an exhaustive parse is possible. Taft (1979, 1994) has offered two somewhat more explicit characterizations of the parsing mechanism. In his earlier work (Taft, 1979; Taft & Forster, 1975), he proposed that decomposition occurs via an affix-stripping process in which the reader parses the affixes from a word using an active, left-to-right search, with the remainder of word used to access the mental lexicon. In a subsequent model (Taft, 1994) he suggested that morphological decomposition could be accomplished within an interactive-activation framework in a network with morpheme detectors situated between letter and word detectors.

Although these hypotheses have proven useful in generating new lines of empirical research, they have not been implemented in any computational models and thus are (like most verbal-description models) rather underspecified. It is fair to wonder how mechanisms of the sort envisioned by these accounts would fare if faced with the challenges that confront readers of natural languages. One such challenge is the presence of pseudomorphological structure: As noted above, a given letter or letter cluster may correspond to a morpheme in some words but not others. Examples in English include not only CORNER and BROTHER, but also BROTHEL and BLUSTER (as well as REAL and BUS). Ambiguity also exists in the other direction—the orthographic realization of a given morpheme may vary across words. The existence of irregular inflections (CAME, GEESE) is one source of this ambiguity; so too are orthographic conventions involving letter doubling (SLIPPER), letter deletion (COMPUTER), and letter change (HAPPINESS). One estimate is that in English, for example, nearly 40% of the morphologically complex words are written such their full form is not simply the concatenation of the intact forms of their constituent morphemes (Baayen et el., 1993). A third challenge for the decomposition process is that the input to the parsing process is surprisingly imprecise. In particular, a wealth of recent evidence has revealed that there is considerable uncertainty in the coding of letter position (as evidenced, for example, by transposed-letter effects—e.g., Andrews, 1996; Perea & Lupker, 2003).

In short, any theory of morphological decomposition must explain how the decomposition process deals with pseudomorphological structure, variability in how morphemes are

realized, and letter position uncertainty. These questions have been the focus of a growing body of recent research (Rastle et al., 2004; Devlin et al., 2004; Diependaele et al, 2005; Feldman et al., 2004; Feldman, O'Connor, & Moscoso del Prado Martín, 2009; Morris et al., 2007; Gold & Rastle, 2007; Marslen-Wilson et al. 2008, McCormick et al., 2008; Taft, 1979; Perea & Carreiras, 2006, Rueckl & Rimzhim, in press; Christianson et al., 2005; Duñabeitia et al., 2007), although further experimental work is clearly needed given some of the empirical conflicts that have yet to be resolved (c.f. Davis & Rastle, 2010; Feldman et al., 2009). Moreover, although this research has spawned a fair amount of discussion, a theoretical account that provides a plausible and detailed answer has yet to emerge.

The failure to address these issues is a shortcoming of the sort of verbal-description models described above, but it is also a shortcoming of the connectionist approach. That is, although models such as Plaut and Gonnerman (2000) and Rueckl and Raveh (1999) have demonstrated networks will 'decompose' a morphologically complex form (in that the organization of its hidden representations reflect morphological structure), these models have not addressed the full range of challenges noted above. Doing so would require solving several of the technical challenges identified in the previous section. To address the problems of pseudomorphological structure and variability in orthographic form, the scale of the model would need to be sufficiently large so that the training corpus would capture the relevant facts about the nature of the language and the writing system. To capture the facts about letter position uncertainty, the input to the network would need to represent the visual properties of a letter string rather than its orthographic form. Based on other simulations (Rueckl & Fang, in preparation), there is reason to expect that under these circumstance letter position would be captured in an appropriately imprecise manner.

Morphology and Phonology

It is striking that although there are many parallels between the literatures on the roles of phonology and morphology in reading, and indeed, although many of the same scientists contribute to both literatures, relatively few studies have investigated whether morphological effects are modulated by phonological variables or vice versa, and more generally, how the morphological and phonological properties of printed words jointly determine the process by which those words are recognized. From almost any theoretical perspective, the independence of the two literatures should be seen as a shortcoming. This is particularly true from the connectionist perspective. One take-home message from the Harm and Seidenberg (2004) simulations is that word recognition involves the cooperative interactions of the O-S and O-P-S pathways. Thus, while models of an isolated O-S pathway (e.g., Plaut & Gonnerman, 2000; Rueckl & Raveh, 1999) can provide important insights about the role of morphology in reading, such models are necessarily oversimplified, particularly given Harm and Seidenberg's demonstration that morphological regularities shape the operation of both the O-S and O-P-S pathways.

This consideration is especially important because, given the relative ease of the O-P mapping: (a) the O-P-S pathway plays an especially prominent role early in the course of reading acquisition (Harm & Seidenberg, 2004); and (b) if the O-P and O-S pathways diverge from a shared hidden layer, the O-P mapping will play a more dominant role in determining the organization of the representations in that layer (Rueckl et al., 1989). These facts suggest that the influence of morphological regularities might be lessened because they occur against the backdrop of stronger phonological influences. On the other hand, the influence of morphology on the O-P mapping could be substantial, particularly with regard to multimorphemic, multisyllabic words.

Gaining clarity on these issues will require both empirical and theoretical developments. On the empirical side, there is a need for more experiments focusing on the interplay of

morphological and phonological influences. Some of these experiments should look at the role of the O-P-S pathway in the computation of word meaning, including whether the influence of morphological regularities can be masked by phonological processes, and conversely whether the O-P-S pathway is itself a source of morphological effects in the computation of meaning. Similarly, it will be important to determine whether morphological regularities shape the computation of phonology, and if so, whether this reflects the direct influence of morphology on the structure of the O-P mapping, or whether instead morphological effects arise in phonological tasks via semantics (i.e., via the O-S-P pathway), in which case these effects would be analogous to the 'Strain effect' (the influence of imageability on word naming, Strain et al., 1995; also see S. Frost et al, 2005).

On the modeling side, the question of how phonological and morphological regularities jointly shape word recognition highlights the need for a large-scale model that incorporates the full triangle model (including the O-S, O-P, and P-S mappings). Simulations of the full model could serve to generate predictions about the outcomes of experiments like those described in the previous paragraph. It is worth noting, though, that such simulations would require solutions to each of the technical issues identified above. The model would have to be of sufficiently large scale. It would require the use of phonological representations appropriate for multisyllabic words and semantic representations that are rich enough to capture the form-meaning correspondences present in a natural language, and it would require an architecture such that the 'orthographic' representations at the front end of the word recognition system are learned and not stipulated by the modeler.

Cross-Language Comparisons

All languages involve morphological processes of one sort or another, and research on morphological processing spans a wide variety of these languages (e.g., Arabic, Basque, English, Finnish, Hebrew, German, Spanish, and so on). Yet, for the most part the study of morphological processing has not taken advantage of the possibilities offered by direct cross-language comparisons. Contrast this with the study of phonological mediation in word recognition, where cross-language comparisons have given rise to major theoretical developments, including the orthographic depth hypothesis (Frost, Katz, & Bentin, 1987) and the grain-size hypothesis (Ziegler & Goswami, 2005).

One notable exception involves the comparison of Hebrew and English. Frost and colleagues (Frost et al., 2000, 2005; Velan & Frost, 2007) have made a strong case that reading in Hebrew and English is accomplished by systems that differ in their organization. In a series of studies they have observed patterns of morphological and orthographic priming effects and letter transposition effects in the two languages. Collectively, these results suggest that compared to English, the representation of Hebrew words is dominated more by morphology (Frost, 2009) and less by orthographic similarity. The Plaut and Gonnerman (2000) simulations described above are consistent with this conclusion. They found that morphological regularities had a greater impact on the hidden representations of the model trained on the morphologically rich language than on the model trained on the impoverished language.

While the Hebrew-English comparison is intriguing, various questions remain open about the underlying basis of these observed differences. First, while the contrast between morphologically 'rich' and 'impoverished' languages has an intuitive appeal, a quantifiable metric characterizing this dimension has yet to emerge. Are rich languages rich because they contain a larger (or smaller) number of morphemes? More systematic and reliable morphological formations? A larger ratio of multimorphemic to monomorphemic words? A larger mean number of morphemes per word? Second, morphological formations in English are largely produced through linear concatenation, whereas a major source of morphological

structure in Hebrew involves the infixing of roots and word patterns. As a consequence, the orthographic structure of Hebrew and English words (in terms of, say, bigram or trigram frequency) may differ in ways that have direct consequences for the organization of the reading system. Relatedly, the characteristics of the mappings from spelling to phonology may also differ systematically as a function of how words are formed in the two languages, and they clearly differ in terms of the kinds of phonological ambiguity associated with written words.

Computational modeling provides one avenue for investigating the potential impact of these various factors on the organization of the reading system. As Plaut and Gonnerman's (2000) simulations demonstrate, modelers have the option of devising artificial languages that vary in precisely controlled ways. Moreover, in addition to observing the effects of these variations on a network's behavior, they can 'peek inside' the network to directly measure the influence of these manipulations on the representations and processing dynamics that give rise to that behavior. Of course, the fact that a model works in a particular manner does not imply that people work in the same way for the same reasons, but modeling results can serve as the basis for generating new empirical predictions.

While this discussion has focused on the contrast between English and Hebrew, it should be clear that the same ideas apply to cross-language comparisons more generally. If something like a 'morphological depth hypothesis' emerges, modeling could serve a useful role in evaluating and refining that hypothesis. But modeling would not be the only direction that cross-language research would need to take. Obviously, the primary ingredient underlying the formulation and evaluation of such a hypothesis would be carefully crafted experimental studies. In addition, a prerequisite for such studies would be the development of metrics characterizing the global ortho-morphological (and ortho-phonological) properties of a given language, as well as the characteristics of the individual words within that language.

Learning

With the exception of connectionist models, virtually all accounts of morphological effects in reading begin with the assumption that readers have representations of morphemes and/or whole words stored in their head, and then ask how these representations must be arranged and operated upon in order to generate the patterns of behavior exhibited in experimental settings. The question of how these representations came to be is generally left for another day or another subfield (e.g., developmental psychology). Happily, however, perhaps this is about to change. Broadly speaking, the relationship between word learning and (written and spoken) word recognition has become the focus of a growing body of research (e.g., Davis & Gaskell, 2009; Leach & Samuel, 2007). With regard to morphology in particular, in a recent paper Rastle and Davis (2008) raised the question of how morphemic representations are acquired.

As Rastle and Davis (2008) note, one important question is whether semantics plays any role at all in the acquisition of morphemic representations. One possibility is that morphemic representations are formed on the basis of purely orthographic information (e.g., by processes that search for either rare or highly frequent letter patterns). In connectionist terms, such a process could be implemented by either unsupervised learning algorithms (e.g., Kohonen, 1982) or by an auto-encoder mechanism that is trained to simply map letter sequences onto themselves (Sibley et al., 2010). However, there is some (albeit limited) evidence that morpheme acquisition is driven specifically by the presence of form-meaning correspondences. For example, Rueckl and Dror (1994) constructed a word learning study in which the systematicity of the mapping from form to meaning was manipulated. For some of the subjects, pseudoword-meaning pairings were constructed so that pseudowords with the same word body were systematically paired with meanings from the same semantic

categories (e.g., *durch-dog, hurch-cat, murch-cow*). For other subjects, the same set of pseudowords and definitions comprised the training set, but the pairings of pseudowords and definitions were constructed so that no such regularities existed (e.g., *durch-dog, hurch-shirt, murch-table*). The results revealed that the structured pairings were easier to learn, and, more importantly, that the pseudowords in this condition were more accurately identified in a tachistoscopic identification task. The latter results suggest that systematic form-meaning correspondences shaped the representations at the 'front end' of the reading system.

The Rueckl and Dror (1994) results are consistent with the kind of learning mechanism embodied in models such as Plaut and Gonnerman (2000) and Rueckl and Raveh (1999). However, one potential concern about the sufficiency of such accounts was raised by Rastle and Davis (2008), who noted that the input representations in both of these models were 'pre-segmented' such that the orthographic realizations of the stems and affixes were represented by different sets of input units. In their view, "this pre-segmented input is what allows these models to recognize orthographic similarity across sets of semantically related words (e.g., distrust, trust, untrustworthy). Thus, some mechanism is required that explains how it might be that form-meaning correspondences drive morphemic segmentation at the orthographic level." (p. 957). The mechanism they proposed, an extension of Bullinaria's (1995, 1997) model, would take as the input a large set of possible orthographic representations of a word. The input representation that most accurately activates the meaning of that word would be given positive feedback, such that over many learning trials the network would converge on input representations that capture morphological structure.

While Rastle and Davis's (2008) proposal represents an important acknowledgment of the need to extend the scope of our models to issues related to learning, their analysis of the earlier models (Plaut & Gonnerman; 2000; Rueckl & Raveh, 1999) is suspect. It is at best an open question whether the ability of these networks to become attuned to morphological regularities was contingent on the use of 'pre-segmented' input representations. In fact, the Harm and Seidenberg (2004) results suggest that this is not the case. In the Harm and Seidenberg model the input units corresponded to letters and the letters that comprised suffixes in some words (e.g., S, D) also appeared in other words as non-morphological final letters (e.g., GAS, WIND).

More broadly, Rastle and Davis's critique of the earlier models highlights the need to distinguish between a theory and its implementation. The design of the input representations is shaped by the modeler's theoretical position in that it is intended to support the kinds of statistical structures that are thought to be relevant to the model's target behavior. Moreover, the properties of these 'stipulated' representations can influence the behavior of a model and significant advances have been made by considering the consequences of different representational schemes (c.f., Plaut et al., 1996, Seidenberg & McClelland, 1989). That being said, the stipulation of a specific representational scheme is also something of a promissory note taken out as an implementational necessity. The 'real' theory (in my opinion, at least) is that representations at every level are learned hidden representations, the organization of which reflects both general computational principles and the structure of the task environment (see Rueckl & Seidenberg, 2009, for further discussion). Of course, the proof is in the pudding, and it this case it remains to be seen if the principles embodied in the triangle model—coupled with an appropriately structured task environment—are sufficient to account for the morphological tuning of the 'orthographic' representations at the front end of the reading system.

Conclusion

Connectionism offers an alternative to more traditional information-processing accounts. In this paper I have attempted to draw out some of these differences with regard to the role of morphological structure in word recognition. At one level these differences concern different characterizations of the processes involved in word recognition and its acquisition. At another level, though, the major difference between these approaches involves their underlying assumptions, the scope of the theory and what counts as an explanation. As noted in the Introduction, in virtually all traditional accounts morphological effects are taken to reflect the structural organization of the lexicon. Thus, behavioral organization is explained in terms of internal organization, and the task of the theoretician is to "reverse-engineer" this internal organization on the basis of patterns of behavior. In the connectionist approach, in contrast, although the effects of morphological structure on behavior are thought to be a consequence of the organization of the hidden representations, this internal organization does not correspond to "the structural properties of the lexicon" in any traditional sense. Instead, the influence of morphological regularities emerges as a consequence of the activation and learning dynamics that control a network's behavior. On this view, internal organization serves a dual role: It provides an explanation for regularities in behavior, but it is also a phenomenon requiring explanation in its own right. The task of the theoretician is to identify the forces that give rise to this organization.

Acknowledgments

References

Andrews S. Phonological recoding: Is the regularity effect consistent? Memory and Cognition. 1982; 10:565–575.

Andrews S. Lexical retrieval and selection processes: Effects of transposed-letter confusability. Journal of Memory and Language. 1996; 35:775–800.

Baayen, RH.; Piepenbrock, R.; van Rijn, H. The CELEX lexical database (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania; 1993.

Barron RW. Word recognition in early reading: A review of the direct and indirect access hypotheses. Cognition. 1986; 24:93–119. [PubMed: 3539489]

Bentin S, Feldman LB. The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. Quarterly Journal of Experimental Psychology. 1990; 42A:693–711. [PubMed: 2287760]

Berko J. The child's learning of English morphology. Word. 1958; 14:150–177.

Besner D, Twilley L, McCann R, Seergobin K. On the connection between connectionism and data: Are a few words necessary? Psychological Review. 1990; 97:432–446.

Brown, R. A First Language: The Early Stages. Harvard University Press; Cambridge, MA: 1973.

Bullinaria JA. Neural network learning from ambiguous training data. Connection Science. 1995; 7:99–122.

Bullinaria JA. Modeling reading, spelling, and past tense learning with artificial neural networks. Brain and Language. 1997; 59:236–266. [PubMed: 9299066]

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

Christianson K, Johnson RL, Rayner K. Letter transpositions within and across morphemes. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2005; 31:1327–1339.

Coltheart M, Curtis B, Atkins E, Hailer M. Models of reading aloud: Dual-route and parallel-distributed-processing approaches. Psychological Review. 1993; 100:589–608.

Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J. DRC: A dual route cascaded model of visual word recognition and reading aloud. Psychological Review. 2001; 108:204–256. [PubMed: 11212628]

Cottrell G, Plunkett K. Acquiring the mapping from meaning to sound. Connection Science. 1994; 6:379–412.

Davis MH, Gaskell MG. A complementary systems account of word learning: neural and behavioural evidence. Philosophical Transactions of the Royal Society B. 2009; 364:3773–3800.

Davis MH, Rastle K. Form and meaning in early morphological processing: Comment on Feldman, O'Connor and Moscoso del Prado Martin. Psychonomic Bulletin & Review. 2010; 17:749–755. [PubMed: 21037177]

Devlin JT, Jamison HL, Gonnerman LM, Matthews PM. The role of the posterior fusiform gyrus in reading. Journal of Cognitive Neuroscience. 2006; 18:911–922. [PubMed: 16839299]

Diependaele K, Sandra D, Grainger J. Masked cross-modal morphological priming: Unraveling morpho-orthographic and morpho-semantic influences in early word recognition. Language and Cognitive Processes. 2005; 20:75–114.

Duñabeitia JA, Perea M, Carreiras M. Do transposed-letter similarity effects occur at a morpheme level? Evidence for morpho-orthographic decomposition. Cognition. 2007; 105:691–703. [PubMed: 17217942]

Elman, JL.; Bates, EA.; Johnson, M.; Karmiloff-Smith, A.; Parisi, D.; Plunkett, K. Rethinking Innateness: A Connectionist Perspective on Development. Cambridge, MA: MIT Press; 1996.

Ervin, SM. Imitation and structural change in children's language. In: Lenneberg, E., editor. New Directions in the Study of Language. MIT Press; 1964.

Feldman LB, O'Connor PA, Moscoso del Prado Martin F. Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. Psychonomic Bulletin & Review. 2009; 16:684–691. [PubMed: 19648453]

Feldman LB, Soltano EG. What morphological priming reveals about morphological processing. Brain & Language. 1999; 68:33–39. [PubMed: 10433736]

Feldman LB, Soltano EG, Pastizzo MJ, Francis SE. What do graded effects of semantic transparency reveal about morphological processing? Brain and Language. 2004; 90:17–30. [PubMed: 15172521]

Forster KL, Davis C, Schoknecht C, Carter R. Masked priming with graphemically related forms: Repetition or partial activation? Quarterly Journal of Experimental Psychology: Human Experimental Psychology. 1987; 39A:211–251.

Frost, R. Reading in Hebrew vs. Reading in English: Is There a Qualitative Difference?. In: Pugh, K.; McCardle, P., editors. How Children Learn To Read: Current Issues and New Directions in the Integration of Cognition, Neurobiology and Genetics of Reading and Dyslexia Research and Practice. Psychology Press; 2009.

Frost R, Ahissar M, Gotesman R, Tayeb S. Are phonological effects fragile? The effect of luminance and exposure duration on form priming and phonological priming. Journal of Memory and Language. 2003; 48:346–378.

Frost R, Deutsch A, Forster KI. Decomposing morphologically complex words in a nonlinear morphology. Journal of Experimental Psychology: Learning, Memory and Cognition. 2000; 36:751–765.

Frost R, Forster KI, Deutsch A. What can we learn from the morphology of Hebrew: A masked priming investigation of morphological representation. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1997; 23:829–856.

Frost R, Katz L, Bentin S. Strategies for visual word recognition and orthographical depth: a multilingual comparison. Journal of Experimental Psychology: Human Perception and Performance. 1987; 13:104–115. [PubMed: 2951484]

Frost R, Kugler T, Deutsch A, Forster K. Orthographic structure versus morphological structure: Principles of lexical organization in a given language. Journal of Experimental Psychology: Learning, Memory and Cognition. 2005; 31:1293–1326.

Frost SJ, Mencl WE, Sandak R, Moore DL, Rueckl J, Katz L, Fulbright RK, Pugh KR. An fMRI study of the trade-off between semantics and phonology in reading aloud. NeuroReport. 2005; 16:621–624. [PubMed: 15812320]

Giraudo H, Grainger J. Priming complex words: Evidence for supralexical representation of morphology. Psychonomic Bulletin and Review. 2001; 8:127–131. [PubMed: 11340857]

Glushko RJ. The organization and activation of orthographic knowledge in reading aloud. Journal of Experimental Psychology: Human Perception and Performance. 1979; 5:674–691.

Gold BT, Rastle K. Neural correlates of morphological decomposition during visual word recognition. Journal of Cognitive Neuroscience. 2007; 19:1983–1993. [PubMed: 17892394]

Gonnerman LM, Seidenberg MS, Andersen ES. Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. Journal of Experimental Psychology: General. 2007; 136:323–345. [PubMed: 17500654]

Hahn, Nakisa. German inflection: Single route or dual route? Cognitive Psychology. 2000; 41:313–360. [PubMed: 11121259]

Harm, MW. Building large scale distributed semantic feature sets with WordNet (Tech. Rep. No. PDP.CNS.02.01). Pittsburgh, PA: Carnegie Mellon University, Center for the Neural Basis of Cognition; 2002.

Harm M, McCandliss BD, Seidenberg MS. Modeling the successes and failures of interventions for disabled readers. Scientific Studies of Reading. 2003; 7:155–182.

Harm MW, Seidenberg MS. Phonology, reading acquisition, and dyslexia: Insights from connectionist models. Psychological Review. 1999; 106:491–528. [PubMed: 10467896]

Harm MW, Seidenberg MS. Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. Psychological Review. 2004; 111:662–720. [PubMed: 15250780]

Joanisse MF, Seidenberg MS. Impairments in verb morphology after brain injury: A connectionist model. Proceedings of the National Academy of Sciences, USA. 1999; 96:7592–7597.

Kello CT, Plaut DC. Strategic control over rate of processing in word reading: A computational investigation. Journal of Memory & Language. 2003; 48:207–232.

Kohonen T. Self-organizing formation of topologically correct feature maps. Biological Cybernetics. 1982; 43:59–69.

Kuczaj SA. The acquisition of regular and irregular past tense forms. Journal of Verbal Learning and Verbal Behavior. 1977; 16:589–600.

Lachter J, Bever TG. The relation between linguistic structure and theories of language learning: A constructive critique of some connectionist learning models. Cognition. 1988; 28:195–247. [PubMed: 3349761]

Landauer TK, Dumais ST. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review. 1997; 104:211–240.

Leach L, Samuel AG. Lexical configuration and lexical engagement: When adults learn new words. Cognitive Psychology. 2007; 55:306–353. [PubMed: 17367775]

MacWhinney B, Leinbach J. Implementations are not conceptualisations: Revising the verb learning model. Cognition. 1991; 29:121–157. [PubMed: 1786671]

Marcus, Brinkman, Clahsen, Wiese, Pinker. German inflection: The exception proves the rule. Cognitive Psychology. 1995; 29:189–256. [PubMed: 8556846]

Marslen-Wilson WD, Bozic M, Randall B. Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. Language and Cognitive Processes. 2008; 23:394–421. [PubMed: 18923643]

Marslen-Wilson WD, Tyler LK. Rules, representations, and the English past tense. Trends in Cognitive Sciences. 1998; 2:428–435. [PubMed: 21227274]

Marslen Wilson WD, Tyler LK, Waksler R, Older L. Morphology and meaning in the English mental lexicon. Psychological Review. 1994; 101:3–33.

McClelland JL, Patterson K. Rules or connections in past-tense inflections: what does the evidence rule out? Trends in Cognitive Sciences. 2002; 6:465–472. [PubMed: 12457897]

McCormick S, Rastle K, Davis M. Is there a 'fete' in 'fetish'? Effects of orthographic opacity on morpho-orthographic segmentation in visual word recognition. Journal of Memory and Language. 2008; 58:307–326.

Michaels CF, Turvey MT. Central sources of visual masking: Indexing structures supporting seeing at a single brief glance. Psychological Research. 1979; 41:1–61.

Morris J, Frank T, Grainger J, Holcomb PJ. Semantic transparency and masked morphological priming: An ERP investigation. Psychophysiology. 2007; 44:506–521. [PubMed: 17498223]

Perea M, Carreiras C. Do transposed-letter effects occur across lexeme boundaries? Psychonomic Bulletin & Review. 2006; 13:418–422. [PubMed: 17048724]

Perea, M.; Lupker, SJ. Transposed-letter confusability effects in masked form priming. In: Kinoshita, S.; Lupker, SJ., editors. Masked priming: State of the art. Hove, UK: Psychology Press; 2003. p. 97-120.

Pinker, Prince. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. Cognition. 1988; 28:73–193. [PubMed: 2450717]

Pinker S, Ullman MT. The past and future of the past tense. Trends in Cognitive Sciences. 2002; 6:456–463. [PubMed: 12457895]

Plaut D. Double dissociation without modularity: Evidence from connectionist neuropsychology. Journal of Clinical and Experimental Neuropsychology. 1995; 17:291–321. [PubMed: 7629273]

Plaut DC, Gonnerman LM. Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? Language and Cognitive Processes. 2000; 15:445–485.

Plaut DC, McClelland JL, Seidenberg M, Patterson KE. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. Psychological Review. 1996; 103:56–115. [PubMed: 8650300]

Plaut DC, Shallice T. Deep dyslexia: A case study of connectionist neuropsychology. Cognitive Neuropsychology. 1993; 10:377–500.

Plunkett K, Marchman VA. U-shaped learning and frequency effects in a multilayered perceptron: Implications for child language acquisition. Cognition. 1991; 38:43–102. [PubMed: 2015756]

Plunkett K, Marchman VA. From rote learning to system building: Acquiring verb morphology in children and connectionist nets. Cognition. 1993; 48(1):21–69. [PubMed: 8403834]

Rastle K, Davis MH. Morphological decomposition based on the analysis of orthography. Language and Cognitive Processes. 2008; 23:942–971.

Rastle K, Davis M, New B. The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. Psychonomic Bulletin and Review. 2004; 11:1090–1098. [PubMed: 15875981]

Rueckl JG, Aicher KA. Are CORNER and BROTHER morphologically complex? Not in the long term. Language and Cognitive Processes. 2008; 23:972–1001.

Rueckl JG, Cave KR, Kosslyn SM. Why are "What" and "Where" processed by two cortical visual systems? A computational investigation. Journal of Cognitive Neuroscience. 1989; 1:171–186.

Rueckl JG, Galantucci B. The locus and time course of long-term morphological priming. Language and Cognitive Processes. 2005; 20:115–138.

Rueckl JG, Mikolinski M, Raveh M, Miner C, Mars F. Morphological priming, fragment completion, and connectionist networks. Journal of Memory and Language. 1997; 36:382–405.

Rueckl JG, Raveh M. The influence of morphological regularities on the dynamics of a connectionist network. Brain and Language. 1999; 68:110–117. [PubMed: 10433747]

Rueckl JG, Rimzhim A. On the interaction of letter transpositions and morphemic boundaries. Language and Cognitive Processes. in press.

Rueckl, JG.; Seidenberg, MS. Computational Modeling and the neural bases of reading and reading disorders. In: Pugh, K.; McCardle, P., editors. How Children Learn To Read: Current Issues and New Directions in the Integration of Cognition, Neurobiology and Genetics of Reading and Dyslexia Research and Practice. New York: Taylor & Francis; 2009. p. 101-134.

Rumelhart, DE.; Hinton, G.; Williams, R. the PDP Research Group. Learning internal representations by error propagation. In: Rumelhart, DE.; McClelland, J., editors. Parallel distributed processing:

Explorations in the microstructure of cognition. Vol. 1: Foundation. Cambridge, MA: MIT Press; 1986. p. 318-362.

Rumelhart, DE.; McClelland, JL. the PDP Research Group. On learning the past tenses of English verbs. In: Mcclelland, JL.; Rumelhart, DE., editors. Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models. Cambridge, MA: MIT Press; 1986. p. 216-271.

Rumelhart, DE.; McClelland, JL. the PDP Research Group. Parallel distributed processing: Explorations in the microstructure of cognition: Volume 1. Foundations. Cambridge, MA: MIT Press; 1986.

Seidenberg MS. The time course of phonological code activation in two writing systems. Cognition. 1985; 19:1–30. [PubMed: 4039640]

Seidenberg MS, Gonnerman LM. Explaining derivational morphology as the convergence of codes. Trends in Cognitive Sciences. 2000; 4:353–361. [PubMed: 10962617]

Seidenberg MS, McClelland JL. A distributed, developmental model of visual word recognition. Psychological Review. 1989; 96:523–568. [PubMed: 2798649]

Seidenberg MS, Petersen A, MacDonald MC, Plaut DC. Pseudohomophone effects and models of word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1996; 22:48–72.

Seidenberg, MS.; Plaut, DC. Progress in understanding word reading: Data fitting versus theory building. In: Andrews, S., editor. From inkmarks to ideas: Current issues in lexical processing. New York: Psychology Press; 2006.

Seidenberg MS, Waters GS, Barnes MA, Tanenhaus MK. When does irregular spelling or pronunciation influence word recognition? Journal of Verbal Learning and Verbal Behavior. 1984; 23:383–404.

Sibley DE, Kello CT, Seidenberg MS. Learning orthographic and phonological representations in models of monosyllabic and bisyllabic naming. European Journal of Cognitive Psychology. 2010; 22:650–668.

Spieler DH, Balota DA. Bringing computational models of word naming down to the item level. Psychological Science. 1997; 8:411–416.

Strain E, Patterson K, Seidenberg MS. Semantic effects in single word naming. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1995; 21:1140–1154.

Taft M. Lexical access via an orthographic code: The Basic Orthographic Syllable Structure (BOSS). Journal of Verbal Learning and Verbal Behavior. 1979; 14:638–647.

Taft M. Interactive-activation as a framework for understanding morphological processing. Language and Cognitive Processes. 1994; 9:271–294.

Taft M, Forster KI. Lexical storage and retrieval of prefixed words. Journal of Verbal Learning and Verbal Behavior. 1975; 14:638–647.

Taraban R, McClelland JL. Conspiracy effects in word recognition. Journal of Memory and Language. 1987; 26:608–631.

Treiman R, Kessler B, Bick S. Influence of consonantal context on the pronunciation of vowels: A comparison of human readers and computational models. Cognition. 2003; 88:49–78. [PubMed: 12711153]

Velan H, Frost R. Cambridge University vs. Hebrew University: The impact of letter transposition on reading English and Hebrew. Psychonomic Bulletin & Review. 2007; 14:913–918. [PubMed: 18087959]

Waters GS, Seidenberg MS. Spelling-sound effects in reading: Time course and decision criteria. Memory and Cognition. 1985; 13:557–572.

Woollams A, Lambon Ralph MA, Plaut DC, Patterson K. SD-squared: On the association between semantic dementia and surface dyslexia. Psychological Review. 2007; 114:316–339. [PubMed: 17500629]

Ziegler JC, Goswami U. Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. Psychological Bulletin. 2005; 131:3–29. [PubMed: 15631549]