1255

ECOLOGICAL PSYCHOLOGY, 14(1–2), 5–19

ARTICLES

The Dynamics of Visual Word Recognition

Jay G. Rueckl
Department of Psychology
University of Connecticut and
Haskins Laboratories

This article provides an overview of a dynamical systems approach to visual word recognition. In this approach, the dynamics of word recognition are characterized in terms of a connectionist network model. According to this model, seeing a word results in changes in the pattern of activation over the nodes in the lexical network such that, over time, the network moves into an attractor state representing the orthographic, phonological, and semantic properties of that word. At a slower timescale, a learning process modifies the strengths of the connections among the nodes in a way that attunes the network to the statistical regularities in its environment. This view of word identification accommodates a wide body of empirical results, a representative sampling of which is discussed here. Finally, the article closes with a discussion of some of the theoretical issues that should be addressed as the dynamical approach continues to develop.

Despite its apparent simplicity, visual word recognition is a remarkable skill. For one thing, the number of words a reader is familiar with is quite large—up to 250,000 for the typical reader. Moreover, word identification involves two distinct tasks. Words have meaning, and words can be pronounced, and a skilled reader gains access to the semantic and phonological properties associated with a written word within a few hundred milliseconds of seeing it. Interestingly, the mappings involved in these two subtasks are quite different. The mapping from spelling to phonology is fairly systematic—words that look alike generally sound alike as well (e.g.,

Requests for reprints should be sent to Jay G. Rueckl, Department of Psychology, Box U-1020, University of Connecticut, Storrs CT 06269. E-mail: jay.rueckl@uconn.edu

lake, take, bake, wake). In contrast, with the exception of morphological relatives (e.g., *bake, baker, baker, bakery*), whether two words look alike has no bearing on whether they are similar in meaning. Yet, despite these differences, the knowledge associated with each mapping can be used to read unfamiliar as well as familiar letter strings. Thus, skilled readers of English can pronounce both *mave* and *zill*, and they can understand what is meant by *greenify* and *ecologize*.

Because of its interesting properties, and because, too, it is easily accessible to experimental investigation, visual word identification has served as an important model system in cognitive psychology for many years. In the late 1800s a sizable literature on word identification blossomed, foreshadowing much of the work that would follow the "cognitive revolution" of the 1950s and 1960s. Since the 1960s visual word identification has been a proving ground for a variety of theoretical constructs, including serial search (Forster & Davis, 1984), spreading activation (McClelland & Rumelhart, 1981; Morton, 1969), and rule-based computation (Coltheart, 1978). Most recently, word identification has served as one of the primary test cases in the exploration of the utility of connectionist networks—and by extension, dynamical systems—as models of cognitive processes. The purpose of this article is to provide an overview of a dynamical systems approach to visual word recognition.

DYNAMICS OF WORD IDENTIFICATION: A CONNECTIONIST APPROACH

A connectionist network is composed of many simple, neuronlike processing units called *nodes*, which 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. A learning algorithm is used to adjust the strengths of the connections (the *weights*) such that the flow of activation is tailored to the structure and task demands of the environment in which the network is embedded (for overviews, see Elman et al., 1996; Rumelhart & McClelland, 1986b).

Many of the early connectionist models (e.g., Rumelhart & McClelland, 1986a; Seidenberg & McClelland, 1989) employed relatively simple feed-forward architectures. In feed-forward networks there is a unidirectional flow of activation from input nodes to output nodes (often with a set of "hidden" nodes in between), and as a consequence these early models had the flavor of quasi-behaviorist stimulus-response associators. It has become increasingly clear, however, that simple feed-forward models are best thought of as simplifications of more powerful (and more interesting) interactive networks. Interactivity (the recurrent flow of activation via feedback connections) allows the pattern of activation in a network to evolve over time, even if the external input to the network remains constant. As a result, interactive networks exhibit *self-organizing attractor dynamics*—over time a network's pattern of activation migrates toward a stable state, and once the network reaches an attractor state it remains there until the input to the network changes.

Many of the attempts to apply the connectionist framework to the understanding of cognitive processes have focused on the case of visual word identification (e.g., Grossberg & Stone, 1986; Harm, 1998; Harm & Seidenberg, 1999; Kawamoto, 1993; Masson, 1995; Plaut, McClelland, Seidenberg, & Patterson, 1996; Rueckl & Raveh, 1999; Seidenberg & McClelland, 1989; Stone & Van Orden, 1994). These efforts have converged on a canonical "triangle model," a network which includes separate layers of nodes responsible for representing the orthographic, phonological, and semantic properties of a word, with hidden units mediating the interactions among these layers. The representations of the triangle model are organized such that similarly spelled words have similar patterns of activation over the orthographic layer, semantically similar words have similar patterns of activation over the semantic layer, and so on. When a given word comes into view, the resulting input initiates a flow of activation within the network. Over time, the network settles into a pattern of activation, which in turn provides input to speech generation and language comprehension processes.

At this point it may be useful to explicitly characterize the structure and operation of the triangle model in the language of nonlinear dynamics. The following points are particularly important:

- The state of a dynamical system is characterized by one or more *state* (or *order*) *parameters*. In the triangle model these parameters are the activation values of the individual nodes. Thus, the network lives in a high-dimensional state space (where the number of dimensions is equal to the number of nodes).
- Dynamical systems are *self-causal:* Changes in the state of a system are a consequence of state-dependent processes. Thus, the behavior of a dynamical system can be characterized by a *flow field* that determines the trajectory of a system through its state space. In the triangle model (as in all recurrent networks), the change in the network's pattern of activation is a function of its current activation, and thus the behavior of the network is state dependent.
- The state space of many dynamical systems includes *fixed points* (attractors and repellers)—states that the system will remain in until it is perturbed. When the triangle model is properly trained, each word has a unique attractor, and the positions of the attractors in the state space are organized to reflect similarities in spelling, pronunciation, and meaning.
- In a dynamical system there are typically one or more *control parameters*, the values of which determine the structure of the flow field (e.g., the location of the fixed points, and so on). The control parameters in the triangle model include the *weights* and the *external input*. The weights are coupling parameters that control the interactions among the nodes and are determined by a learning process that attunes the network to its environment and task demands. The external input is assumed to come from basic visual processes, which feed the network information

about the identity and position of the letters in a written word, and thus has a direct impact on the activation of nodes in the orthographic layer.[1]

Thus, the identification of a written word is a dynamical process that is shaped by two sets of constraints. The *internal constraints* (to borrow a phrase from Gestalt psychology) are embodied in the weights and act to ensure that the states of the components of the network are mutually consistent. The flow field generated by these constraints is multistable, with a large number of attractors (one or more per word) organized to capture the relations among the orthographic, phonological, and semantic properties of the words in the reader's vocabulary. The *external constraints* on the dynamics of word identification reflect the "optical push" that seeing a written word exerts on the lexical system, and they are captured in the triangle model by the external input parameters. If the internal constraints were absent, nodes that receive external input would be driven to a certain level of excitation or inhibition, but nodes that do not receive external input would remain in whatever state they were in. (Thus, in the absence of other forces, the flow field generated by the external input would form an attractor manifold—a hyperplane in the system's state space.) Of course, for the system to work properly, the internal constraints must be present. Hence, in the model the effect on the external constraints is to distort the flow field generated by the internal constraints, strengthening the attractor corresponding to the word now being seen (and perhaps some of its neighbors), and weakening or destroying the attractors corresponding to other words. The direction, speed, and outcome of the system's resulting movement through its state space depends both on the structure of the flow field (as jointly determined by the weights and the external input) and the initial position of the system (its pattern of activation at the time that the external input begins to change).

DYNAMICS OF WORD IDENTIFICATION: REPRESENTATIVE PHENOMENA

The dynamical approach offers a conception of word identification that differs rather dramatically from other kinds of cognitive models. Word identification is not a matter of accessing prestored mental representations, nor does it involve the use of rules to strip a word into its constituents and transform these constituents from one kind of representation to another. Instead, word identification involves self-organizing processes at two timescales. At the faster timescale, commensurate with the rate at which individual words are read, the flow of activation within the lexical network is drawn toward an attractor state, which is jointly determined by

[1]Network models sometimes include additional control parameters (e.g., decay, attentional gain), but in the research of relevance here these parameters have played a relatively minor role and, thus, will not be considered further.

the visual input and the network's pattern of connectivity. The pattern of connectivity is in turn the product of a self-organizing process that occurs at a slower timescale. This slower process adjusts the weights to attune the network to the structure of its environment and task demands.

One way to further clarify the characteristics of these dynamic processes and draw out the differences among the unique aspects of the dynamical perspective is to consider how the dynamical approach accounts for a variety of experimental findings. Of the many sorts of empirical phenomena that are of interest to theories of word identification, only a representative subset will be considered here. Collectively, these phenomena reflect two of the primary characteristics of dynamical systems: the manner in which the dynamics are shaped by the control parameters, and the sensitivity of these dynamics to initial conditions.

Consistency Effects

In a typical word identification experiment, a reader is shown a letter string on a computer screen and asked to either name it aloud or push a button indicating whether or not it is a real word. A starting point for any theory of word identification is the observation that the speed and accuracy with which these responses can be made systematically varies as a function of a number of lexical properties. One such property is word frequency—the more experience a reader has with a word (as estimated by its frequency in a large corpus of text), the faster and more accurate its response in a naming or lexical decision task (Scarborough, Cortese, & Scarborough, 1977). Another, and perhaps less obvious, determinant of reading performance concerns the relation between how a word is spelled and how it is pronounced. Although most words in an alphabetic language adhere to standard spelling–sound correspondences (e.g., *save, mint*), a minority of words violate these correspondences (e.g., *have, pint*). It is commonly observed that words such as *mint* are read faster than words such as *pint*, provided that they are relatively low in frequency. For high-frequency words, effects related to spelling–sound correspondence are generally weak or absent (Coltheart, 1978).

One prominent approach to visual word identification (e.g., Coltheart, 1978) holds that this pattern of behavior reveals the interplay of two kinds of processes: a search process that provides direct access to lexical representations on the basis of a word's orthographic structure, and a process that uses rules to transform a sequence of letters to a phonological code ("sounding out the word") in order to access the lexicon. According to the dual-route account, "regular" words (*mint*) conform to the rules used by the phonological route, but "exception" words (*pint*) do not. As a consequence, although regular words benefit from the operation of both routes, exception words must be read via the direct route. To the degree that the phonological route is involved, it will either slow the correct response or result in an error, and thus regular words are identified faster than exceptions. The interaction of regularity and frequency is attributed to the effect of frequency on the direct route. It is assumed that this route is so efficient in processing high-frequency words that it completes its operation well before the phonological route has a chance to produce an output. Hence, for high-frequency words the routes in effect neither cooperate nor compete, and regular and exception words are read at the same rate.

The dynamical approach offers a different perspective, from which the difference between *mint* and *pint* reflects the statistical structure of the correspondences between spelling and phonology. Most of the English words that end in *int* rhyme with *mint* (e.g., *hint, lint, tint, print*). The one word that is inconsistent with this pattern is *pint*. Thus, for the word *mint, hint* and *tint* are "friends," but *pint* is an "enemy." In terms of spelling–sound correspondences, *mint* is more consistent than *pint*—it has a higher ratio of friends to enemies. From a dynamical perspective, the putative effects of regularity are actually effects of consistency: The more consistent the word, the more quickly and accurately it will be read (Glushko, 1979; Jared, McRae, & Seidenberg, 1990).

In the triangle model, consistency effects are a consequence of the way in which learning structures the pattern of connectivity, and hence the activation dynamics. It is assumed that learning never ceases. Thus, each time a word is read, the weights are adjusted to strengthen that word's attractor. One effect of these weight changes is to improve the network's behavior with regard to that word (i.e., in subsequent encounters with that word, it will be identified more quickly and accurately). In addition, however, because the pattern of connectivity controls the network's response to all of the words that it sees, the learning that results from an encounter with a given word will have consequences for the subsequent processing of other words as well. In other words, learning about one word can interfere with the subsequent processing of another word.

Because they are represented by dissimilar (orthogonal) patterns of activation, interference has minimal effect on words that are dissimilar in spelling, pronunciation, and meaning. For words that are similar along one or more of these dimensions, interference can be either beneficial or detrimental. For example, because *mint* and *hint* are similar in both spelling and pronunciation, the weight changes that strengthen the association between the orthographic and phonological representations of *mint* tend to strengthen the corresponding association for *hint* as well. Thus, learning about *mint* improves the network's performance on *hint*. In contrast, because the bodies of *mint* and *pint* are pronounced differently, the changes in the orthographic–phonological weights resulting from an encounter with the word *mint* are inappropriate for the word *pint*. (In fact, they tend to make the network pronounce *pint* so that it rhymes with *mint*; Seidenberg & McClelland, 1989.)

Consistency effects reflect the cumulative effects of experiences with a word's friends and enemies on the network's pattern of connectivity, and hence on its activation dynamics. The more friends a word has, the faster it will be identified; the more enemies, the slower. Of course, the pattern of connectivity depends not only on the network's experience with a word's friends and enemies but also on its experience with that word itself. As noted above, the effect of spelling–sound consistency on word identification is modulated by frequency—for high-frequency words, consistency has little or no effect. Like readers, the triangle model also exhibits this interaction of frequency and consistency (Plaut et al., 1996; Seidenberg & McClelland, 1989). This interaction comes about because each encounter with a word provides the network with an opportunity to adjust its weights in order to improve its performance. Given enough learning trials, the weights can be adjusted to fully compensate for the interference produced by a word's enemies. Practice makes perfect.

To summarize, frequency and consistency effects reflect the statistical structure of the mappings among the properties of the words in a reader's vocabulary. Over the course of learning, this structure shapes the pattern of connectivity within the lexical system. The pattern of connectivity, in turn, constrains the flow of activation within the system and, hence, its behavior. It is worth noting that this linkage between statistical structure, learning, and behavior provides the basis for understanding a variety of empirical phenomena. For example, readers are influenced by the morphological structure of the words that they read (for reviews, see Feldman, 1995; Henderson, 1985). In the triangle model, morphology plays an important role in word identification because morphology is virtually the only source of structure in the mappings from orthography and phonology to meaning. (That is, with the exception of morphological relatives, words that are similar in form are typically not similar in meaning—contrast *make, made, maker* with *make, take, lake.*) Simulations have demonstrated that networks are sensitive to morphological regularities in much the same way that they are influenced by orthographic–phonological consistency (Plaut & Gonnerman, 2000; Rueckl & Raveh, 1999).

Repetition Priming

The explanation of consistency effects developed above was couched primarily in terms of the effects of learning on the pattern of connectivity in a connectionist network. It may be instructive to restate this explanation using the terminology of nonlinear dynamics. To wit, after a word is identified, the lexical system's control parameters are adjusted so as to strengthen the attractor corresponding to that word. However, changes in the control parameters have global consequences—a change in a control parameter deforms the entire flow field. Consequently, if two words have nearby attractors in both the orthographic and phonological subspaces, strengthening one of their attractors strengthens the other as well. In contrast, if two words with nearby orthographic attractors have relatively distant attractors in the phonological subspace, strengthening one attractor weakens the other. The impact of these transfer effects depends on the frequency with which the unprimed word is encountered. More frequent words have stronger attractors, and thus transfer effects have a negligible impact on them.

Consistency effects reveal the collective influence of many learning events on the dynamics of word identification. However, if each encounter with a word causes a change in the lexical system's control parameters, then in principle it should be possible to observe the behavioral consequences of a single learning event. In fact they are observable—in a phenomenon known as *repetition priming.* Repetition priming is the facilitation in the identification of a word that results from having seen that word recently. Priming effects occur on a scale of minutes, hours, or even days. They can be observed in a variety of experimental tasks, and they influence the behavior of readers at all skill levels (for general reviews, see Roediger & McDermott, 1993; Tenpenny, 1995).

Repetition priming has been modeled in a variety of ways—for example, as a change in a word detector's threshold (Morton, 1969); as a reordering of the lexical entries subjected to a serial search process (Forster & Davis, 1984); and as the influence of episodic memory traces on perception (Jacoby & Dallas, 1981; Kolers, 1979). According to the dynamical approach, priming is a manifestation of the process that adjusts the system's control parameters (i.e., the network's weights) after a word has been identified. These changes strengthen the word's attractor, and as a result, on subsequent encounters with that word the system responds faster and more accurately.

Repetition priming has been extensively studied, both by theorists interested in word identification and theorists concerned with the nature of memory.[2] An extensive review of these findings and their fit with the dynamical approach is provided by Rueckl (in press). For present purposes, it is sufficient to discuss several of the characteristics of repetition priming that are especially illuminating with regard to the nature of the dynamical approach.

One important set of findings concerns transfer effects. Because a change in a system's control parameters deforms its entire flow field, the dynamical account holds that the effects of priming should not be limited to the subsequent identification of that prime, but instead should have consequences for the identification of other words, particularly those that are similar to the prime (and hence have nearby attractors). Consistent with this account, priming has been found to transfer to words that are similar to the prime in spelling, pronunciation, and meaning (for a review, see Rueckl, in press). A particularly robust example of a transfer effect is *morphological priming.* Numerous studies (e.g., Rueckl, Mikolinski, Raveh, Miner, & Mars, 1997; Stanners, Neiser, Hernon, & Hall, 1979) have shown that the identification of a word is facilitated by the prior presentation of a morphological relative (e.g., *walk–walks, compute–computer*). Morphological priming has often been taken to indicate that morphemes are explicitly represented in the mental lexicon. On this view, morphological priming occurs because the identification of morphologically related words involves access to the same mental representations. The dynamical approach offers an alternative account. On this account, morphological priming occurs because morphological relatives are typically similar in spelling, pronunciation, and

[2]In the memory literature repetition priming often goes by the name of *implicit memory.*

meaning. Thus, morphological relatives live close together in state space, so that changes in the control parameters that strengthen the attractor for one word strengthen the attractors for its morphological relatives as well. (Consistent with this account, morphological priming varies with orthographic similarity—e.g., *made* primes *make* more then *bought* primes *buy;* see Rueckl et al., 1997.)

Another interesting aspect of priming from a dynamical perspective is *pseudoword priming.* Pseudowords are pronounceable nonwords such as *mave* and *zill.* An important tenet of the dynamical approach is that words and pseudowords are processed in the same way. That is, seeing a pseudoword, like seeing a real word, causes a flow of activation within the lexical processing network such that over time the network moves into an attractor state, which allows the reader to behave in appropriate ways (e.g., pronouncing the word in a sensible way). Thus, pseudoword identification is an example of *automatic generalization*—unfamiliar inputs are processed in fundamentally the same way as familiar inputs. If this view is correct, then the process that produces repetition priming for words should also produce repetition priming for pseudowords. The empirical evidence supports this claim (for reviews, see Rueckl, in press; Tenpenny, 1995). Although some early findings suggested that repetition priming is purely a lexical phenomenon, it appears that these findings were the consequences of some methodological flaws. The bulk of the evidence overwhelmingly shows that words and pseudowords benefit from repetition priming in much the same way, as the dynamical approach predicts.

Short-Lag Priming

Studies of consistency effects and repetition priming exemplify one strategy for exploring the behavior of a dynamical system: Determine how that system's behavior varies with changes in a control parameter. With regard to consistency effects and repetition priming, the control parameters in question are the lexical network's connection weights.

An equally important investigative strategy it to document how a system's behavior varies as a function of initial conditions. In the case of word identification, this strategy involves examining how the response to a word is conditioned by the state of the system when that word first comes into view. One well-established experimental technique for doing this is the *short-lag priming paradigm.* In this paradigm two stimuli—a prime and a target—are presented in rapid succession. (Typically, the prime is presented somewhere between 25 to 250 msec before the target, although both longer and shorter asynchronies are occasionally used.) The main observation that grows out of this paradigm is that the relation between the prime and target influences how quickly responses to the targets can be made. For example, naming and lexical decision responses are generally faster if the prime and target are related in meaning (e.g., *bird–robin*) than if they are not (e.g., *fruit–robin*). Responses are also influenced by the orthographic, phonological, and morphological relations between the prime and target, although the various forms of priming differ somewhat as a function of prime duration and other task characteristics (for reviews, see Lukatela, Frost, & Turvey, 1999; Neely, 1991).

Superficially, short-lag priming resembles the long-lag repetition effects discussed in the previous section, in that both involve the effects of seeing a word on subsequent processing, albeit at different timescales. However, from a dynamical perspective, these phenomena are fundamentally distinct. Long-term priming is a consequence of learning and changes in the lexical network's pattern of connectivity. In contrast, short-lag priming is an effect of initial conditions. The presentation of the prime causes the system to move toward its corresponding attractor. Because the attractors for similar words are relatively close together, as the system moves toward the attractor for a related prime, it also moves toward the attractor for the target. Thus, because response times increase with the distance between the initial state and the target's attractor, response times are faster if the prime and target are related (for applications of this account to a wide variety of priming phenomena, see Masson, 1995; Plaut & Booth, 2000).

Understood in this way, differences in priming as a function of prime duration and prime–target relation provide a window on the dynamics of word identification. For example, when the prime duration is very short, morphologically related prime–target pairs that differ in semantic relatedness (e.g., *professes–profess* vs. *professor–profess*) produce equivalent levels of priming. However, at longer prime durations priming is correlated with semantic similarity (Raveh, 1999). This pattern is what would be expected on the assumptions that (a) semantic similarity is reflected in the proximity of attractors in the semantic subspace, but not in the orthographic and semantic subspaces; and (b) the external input affects the system's position in the orthographic subspace earlier than it affects its position in the semantic subspace.

Hysteresis

The phenomena discussed thus far (consistency effects, repetition priming, short-lag priming, and so on) have been explored using an arsenal of time-honored and well-honed experimental techniques. Although these techniques provide a glimpse of the dynamical processes underlying word identification, it is also worth noting that none of these methods were specifically developed with dynamical processes in mind. Thus, one of the challenges for the dynamical approach is to develop methodologies that are well suited to elucidating the organization of the dynamics of a cognitive task.

One technique that is often used in the investigation of dynamical systems is to observe how the behavior of a system changes as a control parameter is varied in a smooth and continuous fashion. A well-known success story involving this technique concerns the effects of oscillation frequency (a control parameter) on the dynamics of interlimb coordination. By slowly increasing or decreasing the frequency with which two limbs oscillate, one can observe bifurcations, multistability, hyster-

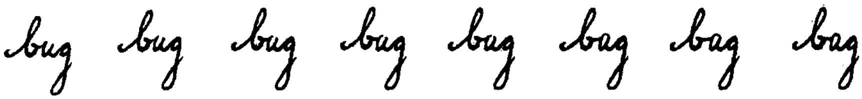

FIGURE 1 An example of a stimulus continuum used in the hysteresis experiments.

esis, and other signatures of dynamical processes. Indeed, the results of studies using this technique have provided a remarkably detailed and elegant understanding of the dynamics of interlimb coordination (see Kelso, 1995).

One attempt to adapt this technique to the study of a more cognitive task was reported by Tuller, Case, Ding, and Kelso (1994) in their study of the dynamics of speech perception. By varying an acoustic parameter that distinguishes the spoken words *say* and *stay* (namely, the duration of the silent interval following the *s*), Tuller et al. constructed a continuum of 20 stimuli that ranged from a clear *say* to a clear *stay*. Listeners heard the tokens of the continuum in order, first in one direction (e.g., *say* to *stay*) and then in the other (e.g., *stay* to *say*). The behavior of the listener's exhibited many of the hallmarks of nonlinear dynamics, including multistability and context-dependent behavior. For example, the listeners often exhibited *hysteresis*—the tendency to remain in the same state, even when other options are available. Thus, a particular token that was heard as *stay* as the continuum changed from *stay* to *say* would be heard as *say* as the continuum changed in the opposite direction.

In a series of unpublished experiments, Jason Fourahar and I applied an analogous technique to the study of written word perception. For example, in one experiment we constructed several continua of handwritten words with two orthographically similar words at their endpoints (see Figure 1). Following Tuller et al. (1994), readers were presented with the tokens of these continua in order, starting at one endpoint, moving through the continuum to the other endpoint, and then back. We observed hysteresis on over 60% of the trials. Moreover, the effects of hysteresis were rather dramatic, often changing the perception of three or four consecutive tokens.

These experiments on hysteresis represent an initial step toward developing methodologies that are particularly appropriate for studying the dynamics of a cognitive task. The articles by Holden and Van Orden in this issue also represent steps in this direction. One of the important challenges for a dynamical approach to word recognition (and other cognitive processes) is to continue to develop experimental methodologies that allow us to ask the kinds of questions that the dynamical approach tells us we should ask.

SUMMARY AND DISCUSSION

The purpose of this article was to provide an overview of a dynamical systems perspective on visual word identification. From this perspective, word identification is a consequence of dynamical processes occurring on (at least) two different timescales. At the faster scale, seeing a word results in changes in the state of the lexical system such that, over time, it moves into an attractor state capturing the orthographic, phonological, and semantic properties associated with that word. At a slower scale, a learning process adjusts the system's control parameters in a way that attunes the system to the structure of its environment and task demands. Connectionist networks provide a framework for modeling these dynamics. Indeed, the account of word identification developed here is a connectionist account, although presented in a way that takes advantage of the rich body of constructs offered by dynamical systems theory.

Over the last decade or so, the dynamical systems perspective (usually couched in terms of connectionist networks) has been applied to a wide variety of empirical phenomena, including not only aspects of the behavior of skilled readers (Harm, 1998; Plaut et al., 1996) but also the effects of brain damage on reading (Plaut et al., 1996) and the problems associated with developmental dyslexia (Harm & Seidenberg, 1999). In the present article a handful of representative phenomena were discussed. These phenomena were chosen because they exemplify central aspects of the dynamical account, including the conceptualization of the connection weights as control parameters, the manner in which the weights are determined by learning and thus come to reflect the structure of the reader's linguistic environment, and the influence of initial conditions on word identification.

Because the application of dynamical systems theory to cognitive processes is relatively new, a brief discussion of several concerns and speculations is in order. One issue of interest to the readers of this journal concerns the points of contact between the present account and the ecological approach to psychology. It seems likely that in comparison to more traditional information processing approaches to word identification, the present account might be more appealing to ecological psychologists in several respects. In particular, both the appeal to self-organizing dynamics and the emphasis on the structure of the environment should seem familiar to proponents of the ecological approach. On the other hand, in the dynamical approach to word identification, relevant environmental regularities are not invariants, but instead are statistical facts concerning the degree of consistency or ambiguity in the mappings among orthography, phonology, and semantics. Moreover, the present approach is representational in the sense that it posits internal states that carry information about an external stimulus, and it is difficult to imagine how an account of a phenomenon such as short-lag priming could do without such states. At the same time, it might be noted that, as Elman (1995) put it, "it is more accurate to think of [the internal state of a network] as the *result* of processing the word, rather than as a representation of the word itself" (p. 207). Perhaps there is a middle ground—a way of characterizing internal states that is rich enough to capture the phenomena of interest to cognitive psychologists without doing violence to the underpinnings of the ecological approach.

Another issue worth considering concerns the dimensionality of the state space occupied by the lexical system. Many of the successes of dynamical systems theory have been achieved by discovering how to characterize the behavior of a complex system in terms of a small number of order parameters. For example, the dynamics of interlimb coordination can be captured by a single order parameter, relative phase (Kelso, 1995). The present account assumes that the lexical system lives in a high-dimensional state space—the dimensionality of a network model is equivalent to the number of nodes that comprise it, and recent versions of the triangle model (e.g., Harm, 1998) include several thousand nodes. Finding a way to dramatically reduce the dimensionality of this space would be attractive on both explanatory and methodological grounds, but at this point it remains unclear whether such a reduction might be achieved. One approach might be to project the high-dimensional system onto a low-order one using principal component analysis or other analytic techniques (e.g., Elman, 1995; Tabor, this issue). Whether this approach could do justice to a reader's ability to distinguish among 250,000 or so different words remains an open question.

Finally, although the dynamical approach puts an emphasis on the statistical structure of the mappings of relevance to reading and how this structure comes to organize the system's flow field, important issues concerning this structure remain unresolved. For example, the measure of orthographic–phonological consistency discussed previously was defined with regard to a word's body (the vowel and final consonants of a monosyllabic word). However, consistency measures could also be defined at other grain sizes (such as individual letters and whole words), and indeed, the consistency of these other units also has behavioral consequences (Cortese & Simpson, 2000; Ziegler, Perry, Jacobs, & Braun, 2001). Although the emphasis on statistical structure seems to be a step in the right direction, at this point we are far from a general theory of how to characterize the behaviorally relevant regularities in the mappings among linguistic properties.

A more fundamental question concerns the origin of these regularities. What simulations of connectionist models have shown is that if a network is exposed to a structured mapping, the regularities in this mapping are captured in the network's pattern of connectivity, and consequently influence the network's activation dynamics. Although this might be the right story for understanding the behavior of a child growing up in a linguistic environment, it is not a sufficient story for understanding the origins of that linguistic environment. The question that lurks is why language has the structure that it does. The hope is that dynamical systems theory can provide the conceptual tools needed to answer this question as well.

ACKNOWLEDGMENT

The preparation of this article was supported by National Institute for Child Health and Development Grant HD–01994 to Haskins Laboratories.

REFERENCES

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). London: Academic.

Cortese, M. J., & Simpson, G. B. (2000). Regularity effects in word naming: What are they? *Memory & Cognition, 28,* 1269–1276.

Elman, J. L. (1995). Language as a dynamical system. In R. F. Port & T. van Gelder (Eds.), *Mind as motion: Explorations in the dynamics of cognition* (pp. 195–223). Cambridge, MA: MIT Press.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

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

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 680–698.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 674–691.

Grossberg, S., & Stone, G. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance, *Psychological Review, 93,* 46–74.

Harm, M. W. (1998). Division of labor in a computational model of visual word recognition. *Dissertation Abstracts International, 60*(02), 0849B.

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

Henderson, L. (1985). Towards a psychology of morphemes. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 15–72). London: Lawrence Erlbaum Associates, Inc.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306–340.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29,* 687–715.

Kawamoto, A. H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language, 32,* 474–516.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior.* Cambridge, MA: MIT Press.

Kolers, P. A. (1979). A pattern-analyzing basis for recognition memory. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing and human memory* (pp. 363–384). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lukatela, G., Frost, S., & Turvey, M. T. (1999). Identity priming in English is compromised by phonological ambiguity. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 775–790.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 3–23.

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

Morton, J. (1969). Interaction of information in word identification. *Psychological Review, 76,* 165–178.

Neely, J. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review, 107,* 786–823.

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

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

Raveh, M. (1999). The contribution of frequency and semantic similarity to morphological processing. *Dissertation Abstracts International, 60*(08), 4280B.

Roediger, H. L., & McDermott, K. B. (1993). Implicit memory in normal human subjects. In H. Spinnler & F. Boller (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63–130). Amsterdam: Elsevier.

Rueckl, J. G. (in press). A connectionist perspective on repetition priming. In J. S. Bowers & C. Marsolek (Eds.), *Rethinking implicit memory.* New York: Oxford University Press.

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

Rueckl, J. G., & Raveh, M. (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain and Language, 68,* 110–117.

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

Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986b). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations.* Cambridge, MA: MIT Press.

Scarborough, D. L., Cortese, C., & Scarborough, H. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 1–17.

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

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

Stone, G. O., & Van Orden, G. C. (1994). Building a resonance framework for recognition using design and system principles. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1248–1268.

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

Tuller, B., Case, P., Ding, M., & Kelso, J. A. S. (1994). The nonlinear dynamics of speech categorization. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1–16.

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science, 12,* 379–384.