In J. Bowers and C. Marsolek (Eds.) (2002).
Rethinking Implicit Memory. Oxford University Press

CHAPTER 4

A CONNECTIONIST PERSPECTIVE ON REPETITION PRIMING

JAY G. RUECKL

The term *repetition priming* refers to a change in the processing of a stimulus as a consequence of a prior encounter with that stimulus. For example, in one of the earliest experimental demonstrations of a priming effect, Neisser (1954) found that studying a word such as PHRASE during a training task lowered the visual duration threshold for that word in a subsequent tachistoscopic identification task. In the half-century since Neisser's seminal study, the literature concerning repetition effects has been dominated by two contrasting theoretical traditions.

One view is that repetition priming is a kind of cognitive 'after-effect'—the manifestation of a temporary change in the process by which words (or other classes of stimuli) are identified. For example, in the classic logogen model (Morton, 1969, 1979), it is assumed that a logogen's threshold (the amount of evidence required to make it 'fire') is lowered each time it fires. Repetition priming is a consequence of this process: If a logogen fires once, it is more likely to fire again due to its lowered threshold. As the logogen model has evolved, it has become more common to speak of the 'flow of activation' rather than the 'accumulation of evidence', and to attribute priming to a persistence in activation rather than a lowering of the threshold. Nonetheless, modern activation theories continue to treat repetition priming as a kind of cognitive after-effect.

An important characteristic of this perspective is that it makes a strong distinction between the mechanism underlying priming and more general learning mechanisms. Thus, for activation theories priming is not a form of learning, and the process by which a representation is primed is unrelated to the process by which that representation came into existence in the first place. In contrast, the other primary approach to repetition priming assumes that it is a manifestation of the very same learning and retrieval processes that give rise to a wide variety of learning and memory phenomena. From this perspective, repetition priming is a form of implicit memory, and as such shares a kinship with phenomena such as classical and operant conditioning and the acquisition of motor skills (Roediger and McDermott, 1993; Schacter, 1987; Squire, 1992).

The activation and memory perspectives represent not only different theories about the processes underlying priming, but also different traditions with respect to theoretical orientation and assumptions about explanatory adequacy. The primary goal of research within the activation tradition has typically been to elucidate the nature of psychological processes within a specific cognitive domain (e.g., the perception of words, faces, or objects). In this tradition priming is as much a diagnostic tool as a phenomenon to be explained. Thus, for example, Morton designed priming experiments to determine whether logogens represent words or morphemes (the latter, according to Murrell and Morton, 1974), and to determine whether spoken and written words are recognized using the same set of logogens (no, according to Morton, 1979).

Research on priming within the memory tradition emphasizes a different set of concerns. From this tradition, the interesting thing about priming lies in its comparison with performance on explicit memory tasks such as recognition and recall. In light of this comparison, repetition priming is relevant to broad issues concerning the organization of memory (e.g., whether there is a single system or separate stores divided along episodic/semantic or declarative/procedural lines), the relationship between encoding and retrieval processes, and the neuropsychological underpinnings of memory functions.

Thus, these two traditions have contrasting strengths and weaknesses. By focusing on specific cognitive domains, activation accounts tend to provide relatively detailed explanations of the mechanisms underlying priming and couch these mechanisms in an explanatory framework that explicitly addresses other classes of phenomena (in the case of word perception, frequency and context effects, short-term priming, and effects of phonological regularity, to name just a few). On the other hand, activation models tend to have the flavor of special-purpose devices whose properties cannot be deduced from or even related to a set of general principles. For example, in the logogen model (Morton, 1969) the threshold-lowering mechanism involves two assumptions: A threshold is initially lowered by a relatively large amount, and over time it returns to a value slightly less than its original value. The first assumption is intended to account for the time course of repetition priming, the second for frequency effects. Neither is logically dependent on the other—the initial change could be small, or the threshold could return to its original value—and if readers turned out to exhibit repetition effects but not frequency effects, the ramifications for the logogen model would be minimal.

Work within the memory tradition has the opposite flavor. The tendency here is to seek explanations that capture the commonalities of priming effects across cognitive domains, and a premium is placed on relating these effects to general, domain-independent principles (e.g., the principle of transfer-appropriate processing, Roediger, 1990, or the consequences of computational incompatibility, Sherry and Schacter, 1987). While this attitude towards explanatory strategy is laudable in many ways, it also has its drawbacks. For example, memory accounts tend to divorce priming effects from other phenomena involving the same perceptual system—if not in principle, at least in practice. (For example, if you want an account of phonological regularity effects within the context of a memory theory, feel free to make one up, but you'll be hard-pressed to find one in the literature.) Moreover, with rare exception (e.g., Hintzman, 1986) theoretical accounts which emphasize broad principles typically have relatively little to say about the characteristics of the mechanisms which implement those principles.

In sum, two distinct traditions have evolved since Neisser's demonstration of repetition priming nearly a half-century ago. One tradition views priming as a form of learning, emphasizes the relation between priming and other memory phenomena, and values explanation in terms of general, domain-independent principles. The other tradition sees priming as a cognitive after-effect, places this effect in the context of other perceptual phenomena, and values explanation in terms of explicitly described, domain-specific mechanisms. To be sure, not every account of repetition priming developed in the last 50 years falls neatly into one or the other of these descriptions. Nonetheless, the contrast drawn above is a reasonable characterization of the literature concerning repetition priming as a whole.

This being said, it should also be noted that the clustering of characteristics which distinguish the two traditions is not a consequence of logical necessity. For example, nothing precludes a perceptual theory from being couched in terms of general principles, nor must theories concerned with memory phenomena eschew attention to processing mechanisms. Thus, while there are historical and pragmatic reasons for the evolution of the theoretical traditions described above, theoretical approaches that break from these traditions remain, in principle, viable options. Indeed, one of the primary motivations for the present volume is to bring together a variety of opinions concerning the benefits (and costs) of alternatives to these traditional perspectives.

One such alternative is offered by the connectionist framework. The purpose of this chapter is to describe the connectionist perspective on repetition priming. The first section provides an overview of the connectionist framework, and more specifically, connectionist models of word perception. The section that follows reviews empirical findings concerning several different issues related to repetition priming—the purpose of this section is both to show how the network approach has served to generate new hypotheses and to demonstrate how this approach allows for new conceptualizations of some old ideas. The chapter closes with a discussion of the relationship between the connectionist approach and accounts which have developed within the activation and memory traditions.

The connectionist approach

From a connectionist perspective, behavior is determined by the massively parallel interactions of many simple, neuron-like processing units. These units, called nodes, communicate by sending excitatory and inhibitory signals to one another, resulting in changes in the pattern of activation across the network. A learning algorithm is used to adjust the strengths of the connections among 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 *et al.*, 1986, among many others.)

A fundamental tenet of the connectionist approach is the principle of *distributed representation*. In connectionist networks, meaning resides in the pattern of activation across a set of nodes, and not in the activation of individual nodes. Thus, representations are distributed across many nodes, and each node plays a role in many different representations. Moreover, because the causal relationships among these representations are encoded by the strengths of the connections, given that each node participates in many representations, each connection must encode information concerning many causal relationships. Hence, distributed representation necessitates *superimpositional storage*, another cornerstone of the connectionist approach.[1]

An important point about distributed representation is that patterns of activation can be more or less similar to one another. Thus, inherent in the notion of distributed representation is the notion of a similarity metric: Some patterns of activation are more similar than others. An upshot of this fact is that a network's behavior follows the *similarity principle*: Similar states tend to have similar consequences. Thus, for example, similar patterns of activation over one set of nodes tend to evoke similar responses over another. Mathematically, the similarity principle follows from the manner in which a node's activation is determined by the excitatory and inhibitory signals it receives from other nodes. Small changes in the pattern of these signals—corresponding to small changes in the pattern of activation of the nodes sending them—generally result in small differences (if any) in the activation of the nodes that receive them. Functionally, the similarity principle reflects the kinds of interference effects that arise as a consequence of superimpositional storage. Because the strength of each connection is determined by the superimposition of knowledge about many different causal relationships, the response to a given input tends to be pulled (for better or for worse) in the direction of the responses associated with other, similar inputs. As we will see, the similarity principle sits behind a wide variety of behavioral phenomena.

In the 'first wave' of connectionist models (e.g., Rumelhart and McClelland, 1986a; Seidenberg and McClelland, 1989), the psychological relevance of distributed representation, superimpositional storage, and the similarity principle was explored using relatively simple feedforward networks. In feedforward networks there is a unidirectional flow of activation from input nodes to output nodes (often with a set of 'hidden' nodes in between). Although feedforward networks have many interesting computational properties, it is now widely appreciated that interactive networks (i.e., networks that include recurrent connections) are much more powerful computationally and much more interesting psychologically. Interactivity allows the pattern of activation in a network to evolve over time, even if the external input to the network remains constant. As a consequence, interactive networks exhibit *self-organizing attractor dynamics*—over time a network's pattern of activation migrates towards a stable state called an 'attractor', and once the network reaches an attractor it remains there until the input to the network changes.

The self-organizing behavior of a connectionist network emerges from the excitatory and inhibitory interactions of its nodes. These interactions depend on the weights of the connections among the nodes, which in turn are determined by an incremental learning process that gradually attunes a network to its environment. Thus, the flow of activation at one time scale is shaped by a learning process that occurs at a slower time scale.

---

[1] It is worth noting that given the prominent role of distributed representation and superimpositional storage in the present account, the term *connectionism* as used here does not extend to models of the 'localist connectionism' variety (e.g., Grainger and Jacobs, 1998; McClelland and Rumelhart, 1981.) In terms of their theoretical assumptions and explanatory strategies, localist connectionist models are fully representative of the activation tradition described above. In many ways the connectionist approach described here should be understood as a rejection of these assumptions and strategies.

Learning influences both the layout of the attractors (i.e., which patterns of activatio stable) and the characteristics of their 'basins of attraction' (e.g., which non-s patterns will result in movement towards a particular attractor, and the speed o movement).

Network models of visual word identification

Many of the attempts to apply the connectionist framework to the understandi cognitive processes have focused on the case of visual word identification (e.g., Gros and Stone, 1986; Harm and Seidenberg, 1999; Kawamoto *et al.*, 1994; Masson, 1995; *et al.*, 1996; Rueckl and Raveh, 1999; Seidenberg and McClelland, 1989; and Stone and Orden, 1994). These efforts have converged on a canonical 'triangle model', a net which includes separate layers of nodes responsible for representing the orthogra phonological, and semantic properties of a word, with hidden units mediating the i actions among these layers. The representations of the triangle model are organized that similarly spelled words have similar patterns of activation over the orthogra layer, semantically similar words have similar patterns of activation over the sem layer, and so on. When a given word comes into view, the resulting input initiates a of activation within the network. The direction, speed, and outcome of this flow de not only on the environmental input, but also on the knowledge of the networ embedded in its weights) and its pattern of activation at the time when the word was encountered.

Advocates of this position maintain not only that the characteristics of a netw behavior resemble those of human readers in important ways, but also that these cha teristics are consequences of distributed representation, superimpositional storage, an similarity principle. For example, both networks (Plaut *et al.*, 1996) and readers (Glus 1979; Jared *et al.*, 1990) exhibit phonological consistency effects: Regular words suc *mint* are read faster than irregular words such as *pint*.[2] For a network, and by theory readers, the phonological consistency effect occurs because the response to a given wo interfered with by the knowledge of other, orthographically similar words. In the cas *mint* and *pint*, *mint* benefits from experience with words such as *lint*, *tint*, and *hint*, whe this knowledge is costly with respect to the identification of *pint*. Seen in this light, phonological consistency effect has the same underlying basis as other ambiguity eff such as the costs associated with the phonological ambiguity of words like *read* and (Seidenberg and McClelland, 1989) and the semantic ambiguity of words such as and *bat* (Kawamoto *et al.*, 1994; but see Joordens and Besner, 1994). All of these phen ena are manifestations of the similarity principle—the tendency for similar states to h similar consequences. This tendency works to a network's advantage in the case of con tent words, but it imposes a cost when it must be overcome, as is the case for inconsist

---

[2] In Glushko's (1979) original formulation of the notion of consistency, a word like *mint* was classified as a (re lar) inconsistent word—only words with bodies always pronounced in the same way (e.g., *pill*) were classi as consistent. However, as the term consistency is now used, words that rhyme with the majority of other wo sharing their bodies (including *mint*) are considered consistent words.

or ambiguous words, as well as, for example, the famous XOR problem (e.g., Minsky and Papert, 1969; Rumelhart *et al.*, 1986).[3]

Networks and repetition priming: a selective review

It is probably fair to say that if repetition effects were not an empirical fact connectionism would not be a viable theoretical option. This is not to say that the study of repetition priming played an important role in the development of the connectionist framework—it didn't. The point is instead that repetition priming so naturally falls out of the behavior of connectionist networks that it would be an embarrassment for the connectionist approach if repetition effects weren't as ubiquitous as they are.

From a connectionist perspective, repetition effects are a consequence of the incremental learning process that continually modifies the connection weights so as to attune the network to the structure of its environment (McClelland and Rumelhart, 1985; Rueckl, 1990; Stark and McClelland, 2000). From this perspective learning never ceases—even the networks underlying highly skilled behavior continue to change with experience (sometimes with rather dramatic results—see Plaut *et al.*, 1996). Repetition effects, like consistency and frequency effects, are a manifestation of this process—but whereas consistency and frequency effects reflect the collective influence of many learning events, repetition effects reveal the influence of individual events on behavior.

Although the link between priming and learning is intrinsic to the connectionist approach, accounts developed within the memory tradition discussed above also make a link between learning and priming. However, the connectionist approach differs from these other accounts in a number of ways: In its commitment to a particular understanding of the learning process (as changes in the patterns of connectivity within a network); In the endorsement of basic principles such as distributed representation, superimpositional memory, and self-organizing dynamics; And in the emphasis on instantiating these principles in computationally explicit, domain-specific processing models that address a variety of empirical phenomena, most of which fall outside the realm of learning and memory as traditionally defined.

These aspects of the connectionist perspective have both theoretical and strategic implications for the investigation of repetition priming. One implication concerns the classification of experimental phenomena: The understanding that priming is a form of learning results in a different partitioning of phenomena than has sometimes been assumed. Another implication is that we should expect the nuances of priming to reflect the operation of principles such as distributed representation and superimpositional memory. A third is that we should strive for an account of priming that links it to other classes of phenomena in deep and insightful ways. In the following sections I illustrate these points by reviewing several lines of investigation into the characteristics of repetition priming.

---

[3] It is worth noting that although these examples illustrate how behavior depends on the degree to which similar *inputs* are mapped onto similar *outputs*, the similarity principle derives from the characteristics of the local interactions between layers of nodes. Thus, although hidden units allow a network to solve the XOR problem (and thus overcome the constraints of the similarity principle), the hidden units are themselves constrained by the similarity principle. The implications of this point are considered below.

Short- versus long-term priming

'Priming' is typically defined as the effect of an encounter with a stimulus on the subsequent processing of another stimulus (typically, a stimulus that is either the same as or related to the prime). On this definition a variety of experimental procedures can be used to study priming. For example, in short-term priming paradigms (e.g., Forster, 1987; Humphreys *et al.*, 1988) the prime precedes the target by a few tens or hundreds of milliseconds, and typically nothing intervenes between the presentation of the prime and target except perhaps a blank screen. On the other hand, in long-term priming paradigms (e.g., Jacoby and Dallas, 1981; Murrell and Morton, 1974) the lag between the prime and target is on the order of seconds, minutes, or even days, and any number of stimuli might occur between the presentation of the prime and target.

A question of theoretical interest is whether the priming effects observed using these two kinds of paradigms are brought about by the same underlying mechanism. According to some activation accounts, the answer is yes. On these accounts, both short- and long-term priming are a consequence of a change in the activation of one or more lexical nodes brought about by the presentation of the prime. However, although such an account is appealing in its parsimony, the empirical facts are that short- and long-term priming are quite different in their characteristics. Short- and long-term priming differ in their persistence (short-term priming is much more limited in duration), in their sensitivity to formal (i.e., orthographic and phonological) and semantic similarity, in the effects of frequency and lexicality, and in their susceptibility to disruption by intervening items (see Forster, 1987; Humphreys *et al.*, 1988; Raveh and Rueckl, 2000, for reviews.) Together, these findings suggest that different mechanisms underlie short- and long-term priming.

In the connectionist framework, short- and long-term priming reflect the operation of two distinct but interrelated processes performed by connectionist networks. As explained above, long-term priming is linked to the learning process that adjusts the pattern of connectivity within the network. In contrast, short-term priming is tied to the flow of activation within the network, a faster process which occurs at a time scale commensurate with the processing of individual stimuli. When a stimulus is encountered, the network moves towards the pattern of activation which represents that stimulus (its 'attractor'). How long it takes to reach the attractor depends not only on its strength (which is determined by factors such as word frequency and phonological consistency), but also on the distance between the attractor and the initial state of the network (i.e., the pattern of activation at the time the stimulus was presented). Short-term priming paradigms involve the manipulation of this initial state: A prime that is similar or identical to the target will tend to put the network closer to the target's attractor than will an unrelated prime. This understanding of short-term priming as, in the vernacular of dynamical systems theory, an 'effect of initial conditions' has been used to account for a variety of findings (see Cree *et al.*, 1999; Kawamoto, 1993; Masson, 1995, 1999; Plaut and Booth 2000; Plaut and Gonnerman, 2000).

The view that short- and long-term priming are the result of different underlying mechanisms is not unique to the connectionist approach (cf. Forster, 1987; Monsell, 1985). Even activation accounts can be formulated such that short- and long-term priming are attributed to different processes. For example, in McClelland and Rumelhart's (1981) interactive activation model, short-term priming might be attributed to the effect of the prime on the target's activation value, whereas long-term priming might be tied to a change in a second parameter associated with each node—the baseline activation value towards which the actual activation eventually decays. This being said, it should also be noted that activation accounts typically treat short- and long-term priming as manifestations of the same underlying process.

Thus, the points to be made concerning the contrast between short- and long-term priming are twofold. First, the connectionist approach attributes these forms of priming to different mechanisms, a view shared by some but not all theoretical perspectives. Second, the connectionist perspective puts the horse before the cart: The existence of each form of priming follows from basic assumptions at the very core of the connectionist framework.

Word and pseudoword priming

Pseudowords are pronounceable nonwords such as *mave* and *zill*. Because pseudowords resemble real words in structure and yet differ from them in familiarity and meaningfulness, the study of pseudoword priming affords an opportunity to disentangle the influence of these variables on repetition priming (and word recognition more broadly). Consequently, pseudoword priming has played an important and somewhat controversial role in the development of theories of repetition priming.

The results of several early studies suggested that, in contrast to the ubiquitous effects of priming on the identification of real words, priming has no effect on the identification of pseudowords. For example, Forbach *et al.* (1974) and Ratcliff *et al.* (1985) observed priming for words but not for pseudowords in the lexical decision task. Similarly, Cermak *et al.* (1985) concluded that, for amnesics, repetition priming facilitates word identification but not pseudoword identification. Diamond and Rozin (1984) reached the same conclusion about amensics' performance on a stem completion task. From the perspective of activation theories, this pattern of results is quite appealing. After all, if repetition priming is brought about by a persisting change in the activation of a lexical representation, then priming should only affect the processing of stimuli that have lexical representations. Pseudowords by definition have no lexical representations, and thus the identification of a pseudoword should not be influenced by repetition.

Upon further examination, however, these early results proved to be somewhat misleading. In particular, pseudoword priming is regularly found in studies employing test tasks other than lexical decision (e.g., Feustel *et al.*, 1983; Rueckl, 1990; Rueckl and Olds, 1993; Bowers, 1996). It seems likely that the failures of Forbach *et al.* (1974) and Ratcliff *et al.* (1985) to observe pseudoword priming are due to the task demands of lexical decision: To the extent that lexical decisions are based on familiarity, any facilitative effects of repetition on the speed of identification may be offset by the tendency for familiarity to inhibit correct 'no' responses (Feustel *et al.*, 1983; Richardson-Klahven and Bjork, 1988). Because repetition has been found to facilitate lexical decisions to pseudowords in some studies (e.g., Danenbring and Briand, 1982; Kirsner and Smith, 1974; Scarborough *et al.*, 1977) and inhibit it in others (e.g., Bowers, 1994), it seems likely that the trade-off between these factors may be modulated by methodological variables.

In contrast to the results of Diamond and Rozin (1984) and Cermak *et al.* (19
amnesic readers have also been found to exhibit pseudoword repetition effects (Haist *et*
1991; Keane *et al.*, 1991; Musen and Squire, 1991). In retrospect, the conclusions dr
from the earlier studies were based on rather unconvincing evidence. In the Diamond
Rozin (1984) study, priming was measured by performance on a stem completion task
rather odd task ('Name a nonword beginning with zil___') that seems quite likely to inv
explicit memory strategies. In the Cermak *et al.* (1985) study, the pseudoword prim
effect for amensics was not significant, but numerically it was nearly twice as large as
comparable (and significant) effect on word identification.

Thus, the empirical facts are clear: Repetition facilitates the identification of both wo
and pseudowords. The theoretical question raised by these facts is whether word
pseudoword priming have a common basis. For activation accounts, the answer must cle
be 'no'. If word priming is due to changes in the activation of pre-existing lexical represen
tions, then a different mechanism must be responsible for pseudoword priming. Thus,
example, word and pseudoword priming have been attributed to processes involving 'a
vation' and 'elaboration' (Dorfman, 1994) or 'lexical' and 'episodic' memory (Feustel *et*
1983). An appealing aspect of these accounts is that the process underlying pseudow
effects might open the door to a theory of how lexical representations come into existen
and thus positing such a process could be justified on a priori grounds. On the other ha
resort to a second process can also be seen as the application of a theoretical strate
employed all too often in cognitive psychology: For each effect posit a new cause.

The connectionist approach joins with certain other accounts (e.g., Marsolek *et al.*, 19
Keane *et al.*, 2000; Schacter, 1992; Whittlesea, 1987) in rejecting the claim that word a
pseudoword priming are brought about by different mechanisms. A general tenet of t
connectionist approach is that both familiar and unfamiliar inputs are processed in the sa
way. Seeing a pseudoword, like seeing a real word, causes a flow of activation within the le
cal processing network. Over time the network moves into an attractor state that allows t
reader to behave in appropriate ways (e.g., pronouncing the word in a sensible way). Th
pseudoword identification is an example of *automatic generalization*—unfamiliar inputs a
processed in fundamentally the same way as familiar inputs. Due to the similarity princip
(the tendency for similar states to have similar consequences), the response to an unfami
iar input resembles the responses associated with similar previously experienced inputs.

On this view, just as words and pseudowords are identified in the same way, so too are wor
and pseudoword priming the consequence of the same process—the learning algorithm th
adjusts the network's pattern of connectivity. The goal of the learning process is to adapt th
network to its environment by strengthening an input's attractor (and repositioning it if neces
sary) so that in future encounters with that input the network responds quickly and accuratel
Because the environment may change over time, and because the network's ability to sampl
the environment is limited, the attunement of the network to its environment must be a
ongoing process. Thus, each event provides an opportunity for learning, whether that event i
the first or the millionth encounter with a stimulus, and each opportunity is taken. Word
priming and pseudoword priming are manifestations of the same underlying process.

Several general points can be made about this perspective on pseudoword priming. First,
whereas the connectionist account distinguishes between short- and long-term priming,

attributing each class of phenomena to different underlying mechanisms, it treats word and pseudoword priming as fundamentally the same. In its stance on each of these contrasts, the connectionist approach differs from some accounts and resembles others.

Second, the connectionist account of pseudoword priming follows directly from the foundational principles of the connectionist framework. Because networks employ distributed representations, there is no need to draw a qualitative distinction between the representations of words and pseudowords: Both familiar and unfamiliar inputs are represented by distributed patterns of activation which are constructed 'on the fly' via self-organizing activation dynamics. Because distributed representation is coupled with superimpositional memory, knowledge acquired from past experiences with real words and distributed across the network's connections can be used to respond appropriately to pseudowords as well as real words. Finally, because learning is a process of attunement rather than a process designed to create lexical representations and store them in memory, it must be understood as an ongoing process that results in changes in the pattern of connectivity each time an input is processed, regardless of the lexical status or familiarity of the input.

A third general point to be made here is that the contrast between the connectionist and two-process accounts of priming parallels several prominent debates concerning other aspects of word perception and production. One of these debates concerns reading aloud. According to dual-route models (e.g., Coltheart, 1978; Coltheart *et al.*, 2001), readers have two options: Words can be identified via a lexical look-up route or via the application of grapheme-phoneme conversion rules. Both routes are thought to be necessary because pseudowords cannot be read via the lexical route and rule-violating exception words (e.g., *pint, have*) cannot be read correctly using sublexical rules. (Regular [rule-following] words can be read via either route.) In contrast, connectionist accounts (Plaut *et al.*, 1996; Seidenberg *et al.*, 1994; Seidenberg and McClelland, 1989) reject the claim that different processes are needed to name exception words and pseudowords. According to these accounts, regular words, exception words, and pseudowords can all be named by a single mechanism—a network trained to map orthographic representations to phonological representations.[4]

A similar (and equally hotly contested) debate concerns the production of English past-tense forms (cf. Rumelhart and McClelland, 1986b; Pinker and Prince, 1988). Here too, dual-route accounts hold that skilled performance depends on two processes: a rule-based mechanism that can be used to generate the past-tense forms of regular words and pseudowords (e.g., *wug–wugged*), and a second process (lexical look-up) that is needed for rule-violating exception words (e.g., *ran, spoke*). Again, from a connectionist perspective a single mechanism suffices—a network capable of producing the past-tense form for regular words and pseudowords is also capable of generating the past-tense forms of exception words.

Although the phenomena involved are rather different (repetition priming, word naming, past-tense production), in all three cases the connectionist position is founded on the same basic assumptions (distributed representation; superimpositional storage; incremental learning) and prejudices (that the same fundamental principles are at work across cognitive domains; that a theory of skilled behavior should not be divorced from a theory of learning; that often apparently distinct phenomena turn out to be different facets of the same underlying mechanism). It is also worth noting that in all three cases the two-process accounts are forced into their positions by the discrete nature of the underlying representations and processes. A stimulus either has a lexical representation or not; it either follows the rules or not. The use of distributed representation and superimpositional memory obviate the need for a language user (or a model thereof) to draw such distinctions. Effects of lexicality and regularity are quantitative, not qualitative.

Interference and similarity effects

According to the connectionist account, when a reader encounters a word, the resulting weight changes are distributed across the lexical network's entire connection matrix. Because these weight changes act to strengthen a word's attractor, the next time that word is seen it will be responded to faster and more accurately. The effects of a learning event are more far-reaching than this, however. In particular, because networks employ distributed representation and superimpositional storage, the response to every word is determined by the same set of connections. Thus, learning about one word has consequences for the identification of other words as well. These consequences are reflected in the similarity principle—the tendency for similar inputs to be mapped onto similar outputs. Depending on the circumstances, this tendency may be either beneficial or detrimental.

To further illustrate this point, consider a gravitational system such as a solar system comprised of numerous massive bodies. Each body exerts a gravitational force which attracts nearby objects, but in general the gravitational field controlling the movement of an object is determined jointly by all of the bodies within the system. Thus, changing the mass of one body, or introducing a new body such as a passing star, has consequences for the gravitational field everywhere within the system (although they may be negligible at great enough distances). Note too that the interaction of the gravitational forces generated by different bodies depends on their spatial relationships: When the moon is full, the earth and sun pull in the same direction; when the moon is new they tug in opposite directions.

Gravitational systems provide a useful metaphor for understanding the dynamics of connectionist networks (Tabor and Tanenhaus, 1999). The patterns of activation acting as attractors in the state space of a network are analogous to the bodies of mass acting as attractors in the physical space of a solar system, with similar words occupying nearby positions in space. Similarly, the trajectory of the network's movement through its state space (the changes in its pattern of activation over time) corresponds to the trajectory of an object through the solar system. When learning strengthens a word's attractor, the consequences are (due to superimpositional memory) global: The entire activation flow field changes, although the impact of the changes diminishes with the distance from that word's attractor. Finally, how learning about one word affects the apparent strength of the attractor of another word is situation-dependent: Aligned in one way, strengthening one word's attractor has the effect of moving the system more quickly towards a neighboring word's attractor as well;

---

[4] Because phonological codes can be computed either directly or via semantics, there is a sense in which the triangle model is a dual-route model. However, in the triangle model the two routes are more redundant rather than complementary and both routes operate in accordance with the same principles. Moreover, although the relative contributions of the two routes may differ for different types of stimuli, in the triangle model the phonological route plays a fundamental role in the naming of regular words, pseudowords, AND exception words (Harm, 1998; Plaut, 1997; Plaut *et al.*, 1996).

aligned in a different way, the effect may be to slow down the rate with which the system moves towards the neighbor, or perhaps prevent the system from getting there at all. In either case, the degree to which learning about one input influences the response to another will vary with similarity—the greater the similarity, the more the interference.

The susceptibility of networks to interference is well-documented, and the factors that modulate its impact are well-understood (cf. Hetherington and Seidenberg, 1989; McClelland *et al.*, 1995; McCloskey and Cohen, 1989; Murre, 1992; Ratcliff, 1990; Rueckl, 1993). What is perhaps less well-appreciated is that from a connectionist perspective interference—the effect of learning about one input on the response to other inputs—underlies a variety of seemingly disparate phenomena. For example, interference clearly underlies forgetting in the A–B A–C paired-associate task (McCloskey and Cohen, 1989). However, from the right vantage point interference also gives rise to the ability to pronounce a nonword aloud or generate its past-tense form. In all three cases, the response to one input is influenced by the weight changes made during learning events associated with other inputs. The fact that transfer is beneficial in some cases and detrimental is not uninteresting, but it should not obscure the fact that positive and negative transfer are both forms of interference.

Thus, although it may seem odd to think of nonword naming and forgetting in the paired-associate task as related, from the connectionist perspective they are both consequences of superimpositional storage. To take another example, consider the consistency effect in reading aloud described in a previous section: Words with many friends (words with bodies that are pronounced the same way in many words, such as the *-int* in *lint, hint, tint,* and *mint*) are named faster than words with many enemies (words with the same body but a different pronunciation; *hint, tint,* and *mint* are enemies of *pint*). Because friends are similar in both spelling and pronunciation, they require similar patterns of connectivity in the weights between these layers. Thus, the weight changes made to improve the response to one word will generally improve the response to its friends as well. In contrast, because enemies are similar orthographically but not phonologically, they require different patterns of connectivity. In this case, learning about one word weakens the response to its enemies. Consistency effects reflect the degree to which the response to a word benefits from knowledge of its friends or suffers from the knowledge of its enemies.

To summarize, because networks employ superimpositional memory, strengthening the attractor for one word has consequences for the identification of other words as well. The consistency effect provides one demonstration of the effects of this cross-talk on word identification, but the account predicts that interference should show up in other guises as well. In particular, it predicts that because repetition priming is a manifestation of learning, priming one word should influence not only the subsequent identification of that word, but also the identification of other words, especially words that are similar to the prime. The following sections review the evidence related to this prediction, beginning with evidence of positive transfer and then turning to the negative effects of interference.

Positive transfer effects

One of the first studies to look for evidence of transfer effects in repetition priming was the seminal experiment by Murrell and Morton (1974). Subjects in their experiment were asked to identify tachistoscopically presented words that differed in their relationships to the words on a previously studied list. Compared to words that were unrelated to any of the study items, words that had been presented on the study list were more accurately identified, as were words that had been primed by a morphological relative (e.g., *cars–car*). In contrast, priming had no effect on the identification of target words that were related to study items orthographically but not morphologically (e.g., *card–car*). Thus, Murrell and Morton found evidence for a transfer effect in long-term repetition priming, but in their study transfer extended only to morphologically related words.

The implications of morphological priming will be considered in a subsequent section. For now the question we will address is whether transfer effects occur on the basis of orthographic similarity. The Murrell and Morton (1974) results suggest not: The identification of a target word was neither facilitated nor impaired as a result of studying an orthographically related word. A variety of other studies have yielded similar patterns of results (e.g., Napps and Fowler, 1987; Ross *et al.*, 1956; Ratcliff and McKoon, 1996; Rueckl *et al.*, 1997). However, before the conclusion is drawn that repetition priming has *no* effect on orthographically similar words, it should be noted that the results of several other studies suggest exactly the opposite.

For example, evidence of a purely orthographic transfer effect in word identification was reported by Feustel *et al.* (1983), who found that the naming latency for a letter string emerging from a background noise mask was faster when it had been preceded by an orthographically similar item. For words, this effect was statistically significant only when the prime had already been presented three or four times: after one or two presentations of the prime, a non-significant trend was obtained. In contrast, with pseudoword targets a single presentation of the prime was enough to produce similarity priming, although again the amount of facilitation increased with the number of presentations of the prime.

Transfer effects based on orthographic similarity have also been found when target items were primed by clusters of orthographically similar items. In one line of experiments, this was accomplished by constructing all of the primes from a set of 13 letters. In experiments employing this *similarity priming* paradigm (Rueckl, 1990; Rueckl, submitted; Rueckl and Olds, 1993), the identification of new items that were similar in spelling to the study items (by virtue of being composed of letters from the same letter set) were more accurately identified than items that were composed only of letters not appearing during the study task. Similarity priming was observed for both words and pseudowords, and cross-experiment comparisons suggested that the transfer effect was due to facilitation in the identification of the similar items, rather than a decrement in the identification of the dissimilar items. An effect of positive transfer on pseudoword identification was also reported by Whittlesea (1987), who generated cohorts of orthographically similar pseudowords by adapting a technique commonly used in categorization studies: Several five-letter pseudowords were chosen as 'prototypes', and a set of 'category members' were generated by replacing one, two, three, or four letters in the prototype. Whittlesea found that the identification of a previously unseen pseudoword was facilitated by the prior presentation of pseudowords from the same category; the degree to which a new pseudoword benefited from transfer increased with its similarity to the primes (and hence, typically, with its similarity to the prototype).

Facilitative effects have also been found with regard to other dimensions of similarity. Consider, for example, the results of a series of experiments investigating the effects of rhyme priming on visual stem completion. During the study phase of these experiments, subjects read either a list of words (Mandler *et al.*, 1986; Mandler *et al.*, 1990) or a poem (Overson and Mandler, 1987) that contained words that rhymed with a target word, but not the target word itself. Relative to an unprimed baseline condition, rhyme priming both increased the likelihood that target words were generated during a subsequent stem completion task (Mandler *et al.*, 1986, 1990; Overson and Mandler, 1987) and reduced the time needed to generate these completions (Mandler *et al.*, 1990). Although the effect of rhyme priming appears to be fairly short-lived relative to full identity priming (Mandler *et al.*, 1986), the fact that it occurs at all provides evidence of similarity-based transfer effects in repetition priming.

Semantic similarity can also give rise to transfer effects. For example, although it is fairly widely believed that long-term priming does not occur on the basis of semantic relatedness, Becker *et al.* (1997) observed just such an effect using a variant of the similarity priming paradigm. In their experiment, lexical decisions were facilitated for target words that were semantically related to a number of words on a study list, even though the critical targets were not themselves presented during study. A similar result was obtained by McDermott (1997) using Deese's (1959) paradigm, which also involves priming a target word with a set of semantically associated words. Finally, Joordens and Becker (1997) demonstrated that under some conditions the presentation of even a single semantic associate can give rise to similarity priming.

In sum, a number of results suggest that the processing of a word or pseudoword can benefit from a recent experience with a similar item. However, it must also be noted that a number of studies looking for such effects have failed to find them. To some extent, methodological factors are probably responsible for this 'now you see it, now you don't' pattern. In most of the studies that failed to detect positive transfer, a target was primed by a single similar item (Murrell and Morton, 1974; Napps and Fowler, 1987; Ratcliff and McKoon, 1996; Rueckl and Mathew, 1999). In contrast, many of the studies where positive transfer was observed involved either multiple related primes (Becker *et al.*, 1997; McDermott, 1997; Rueckl, 1990; Rueckl, submitted; Rueckl and Olds, 1993) or multiple repetitions of a single related prime (Feustel *et al.*, 1983; also see Seidenberg and McClelland, 1989, Fig. 15). Moreover, positive transfer effects in priming appear to be stronger for pseudowords than for real words (Feustel *et al.*, 1983; Rueckl, 1990), and stronger for 3rd graders than for skilled adult readers (cf. Feldman *et al.*, in press; Rueckl *et al.*, 1997). Together, these results suggest that positive transfer effects are (not surprisingly) relatively weak, and hence are more likely to be detected with items that benefit from multiple learning events or have relatively weak attractors to begin with.

Negative transfer effects

Another likely reason for the somewhat elusive nature of positive transfer effects is that priming one word can also have detrimental consequences for the identification of other words. Thus, whether positive or negative transfer will be observed depends on the balance between these cooperative and competitive effects.

Investigations of negative transfer effects have taken two general approaches. One approach seeks to determine whether priming one word reduces the magnitude of repetition effects for other primed words. For example, Jacoby (1983) had subjects perform both a study task (naming words aloud) and a test task (tachistoscopic identification) on each of five successive days. As expected, previously presented words were more accurately identified than unstudied words. However, the magnitude of this effect did not diminish over the five days, thus providing no evidence for the build-up of proactive interference. Moreover, a final identification task at the end of the fifth day revealed approximately equivalent levels of priming for words studied on days 1–5, thus suggesting that priming had not been diminished by retroactive interference. Experiments by Graf and Schacter (1987) and Sloman *et al.* (1988) also failed to find any evidence that the priming of one word has retroactive or proactive effects on the priming of another word.

Although the results described above suggest that repetition priming is not susceptible to negative transfer, the findings of several other studies point towards the opposite conclusion. Mayes *et al.* (1987) had subjects study two lists of semantically related word pairs. The word pairs for the second list were constructed by pairing each cue from the first list with another related word (thus, if *bee–wasp* was on the first list, *bee–honey* was on the second). Mayes *et al.* found more priming on a free association task for words from the first list than for words from the second list, suggesting that proactive interference had reduced the priming effect in the second phase. Similarly, Nelson *et al.* (1989) found evidence of retroactive interference in an experiment investigating priming in the word fragment completion task. Nelson *et al.* found that subjects who saw two lists were less likely to complete a fragment with a word from the first list than were subjects that only saw that list, provided that the words on the two lists were orthographically related.

Although the Mayes *et al.* (1987) and Nelson *et al.* (1989) studies appear to provide evidence of both proactive and retroactive interference in implicit memory, both studies employed 'non-unique cues'—cues for which more than one primed word was an acceptable response. This raises the possibility that the reductions in priming observed in these studies were artifacts arising from the constraint that only one primed item could be generated on each trial. However, Booker (1992) reported a series of experiments in which statistical procedures were used to correct for this response constraint and still obtained evidence that priming was reduced by proactive interference. Rueckl (submitted) also found that priming was reduced by proactive interference in a variant of the similarity priming paradigm where the presentation of primes composed of letters from one 13-letter set was followed by the presentation of primes composed of letters from the complementary 13-letter set.

Thus, one line of evidence suggests that priming one word reduces the magnitude of repetition effects for other primed words. A second line of evidence indicates that negative transfer influences the processing of unprimed words as well. For example, in the same experiments that demonstrated proactive interference on identity priming, Rueckl (submitted) also showed that positive transfer effects were eliminated when a set of primes composed of one set of letters was followed by another set of primes composed of different letters. Thus, for the unprimed test words, what one set of primes giveth, another set of primes taketh way.

Negative transfer effects on unprimed words have been observed in other experimental paradigms as well. For example, Smith and Tindell (1997) found that subjects were less likely to find the target word completing a word fragment (e.g., *symphony* for *s_mp_o_y*) if they had been primed with an orthographically similar word (e.g., *sympathy*). (Also see Lustig and Hasher, 2001.) Bowers *et al.* (in press) reported an experiment where positive transfer was found for word pairs that rhymed (e.g., *boast–toast*), but negative transfer was found for word pairs that did not (e.g., *pint–mint*).

One paradigm that has played an especially prominent role in the study of negative transfer effects is the two-alternative forced-choice task (Ratcliff and McKoon, 1996, 1997). In this paradigm the presentation of a tachistoscopically presented target word is followed by the presentation of two response alternatives; the reader's task is to decide which of the alternatives had been presented as the target. Studies using this paradigm have consistently found that relative to an unprimed baseline, performance is better if the target had been presented during an earlier task, but worse if that non-target alternative had been primed (e.g., Bowers, 1999; Ratcliff and McKoon, 1996, 1997; and many others). Moreover, costs associated with a primed non-target alternative occur when the alternatives are orthographically similar (e.g., *lied–died*) but not when they are dissimilar (e.g., *sofa–died*). Although certain issues concerning the forced-choice paradigm remain rather controversial,[5] the resulting body of findings provides clear evidence of negative transfer: Under some conditions priming one word diminishes the likelihood that another word will be correctly identified.

Interference and similarity effects: redux

In summary, the evidence concerning similarity-based transfer effects in priming is mixed in two respects. First, transfer effects based on similarity in form or meaning are somewhat elusive—they've been found in some experiments, but not in others. Second, when they do occur, their effect is sometimes beneficial and sometimes detrimental.

The elusive nature of these effects suggests that transfer effects are relatively small, especially for skilled readers reading familiar words, and thus that various methodological factors can either bring them into view or submerge them into the background noise. A number of likely factors have been identified, including the number and variety of the primes that are similar to a given target, the familiarity of that target, and the amount of experience that a reader brings into the experimental setting.

The fact that transfer effects can take the form of either costs or benefits probably reflects several trade-offs among the consequences of the learning that takes place when a word is identified. First, at the microlevel modifying the weights to strengthen the attractor for one word will generally have both helpful and costly effects on the degree to which those weights are also appropriate for a similar word. For example, learning about *mint* increases the strength of the association between the word body *-int* and the word rime */-int/*, thus benefiting processing of *hint* as well as *mint*. On the other hand, the association between *-int* and the phonological onset */m/* also get stronger, to the benefit of *mint* but at a cost to *hint*.

---

[5] Controversies concerning the forced-choice paradigm include whether the effects of priming reflect increased sensitivity as well as bias (cf. Bowers, 1999; Neaderhiser and Church, 2000; Ratcliff and McKoon, 1997), and the extent to which the costs and benefits observed in this paradigm reflect the effects of priming on the processing of the target or on the processing of the response alternatives (see Bowers, 1999, for discussion).

A second potential trade-off concerns the competitive interactions among attractors in determining the activation dynamics. Although priming may increase the strength of the attractor for a word that is similar to the prime, the attractor for the prime itself will be strengthened to an even greater extent. Thus, in situations where these attractors are put into competition, the effect of priming is to put the attractor for the similar word in a competitive disadvantage, even if the attractor for the similar word is stronger in an absolute sense (e.g., in terms of *energy*; Hopfield, 1982; Masson, 1995). The analogy with a gravitational system illustrates this point. Given one alignment (sun, earth, moon) increasing the mass of the sun pulls the moon towards the earth; given another alignment (sun, moon, earth), increasing the mass of the sun acts to pull the moon away from the earth. In this case, the border between the basins of attraction for the sun and the earth has shifted towards the earth.[6]

The gravitational analogy suggests that negative transfer effects should be most likely to occur in tasks that highlight the competition among similar words. Thus, perhaps it is not surprising that some of the strongest evidence for negative transfer has comes from experiments using the two-alternative forced-choice procedure. By specifying two response alternatives, there is a sense in which this task seems to pit their attractors in a direct competition, with priming tilting the outcome towards the repeated alternative. That negative transfer doesn't occur when the alternatives are dissimilar (e.g., *lied–sofa*) suggests that little competition occurs between distant attractors.[7]

Although this line of reasoning seems promising, much work remains before its promise can be fully evaluated. In particular, the methodological factors that determine whether similar words will cooperate or compete are not well understood. In this light, it is worth noting that a similar issue has cropped up with regard to neighborhood effects in word identification. Neighborhood effects (effects of the number of words that are similar to a target, as well as the frequency and distribution of these 'neighbors') have been a focus of intense scrutiny in the word identification literature, and for reasons not yet fully understood, they turn out to appear in both facilitative and inhibitory forms (see Andrews, 1997; Grainger and Jacobs, 1996, for reviews). Because both neighborhood effects and transfer effects in priming involve the mechanisms by which the identification of one word is influenced by experience with other words, it is not unreasonable to speculate that identifying the factors that determine the balance between competition and cooperation in one domain will help clarify the results of the other as well.

Morphological priming

As noted above, Murrell and Morton (1974) found that the identification of a target word (e.g., *car*) was facilitated by the prior presentation of a morphologically related word

---

[6] To make the analogy more complete, one could suppose that when the mass of the sun in increased, the mass of the earth is also increased, but to a lesser extent. The primary point about competition and cooperation holds in either case, however.

[7] These intuitions are incorporated into Ratcliff and McKoon's (1997) counter model, which was devised largely to account for results involving the forced-choice task. One might argue that to the degree that the counter model succeeds, it does so because the processes it proposes mirror the activation and learning dynamics of a connectionist network.

(e.g., *cars*). In the years since, this basic finding has been replicated and extended in a variety of ways (see Feldman, 1991; Henderson, 1985, for reviews). Morphological priming has been demonstrated in a variety of languages (including Hebrew, Bentin and Feldman, 1990, Serbo-Croatian, Feldman and Fowler, 1987, and Italian, Burani and Carramazza, 1987, to name just a few). Moreover, although morphologically related words are usually related in form (i.e., spelling and pronunciation) and meaning, morphological priming cannot be attributed solely to similarity along these dimensions (Bentin and Feldman, 1990; Napps, 1989; Napps and Fowler, 1987; Stolz and Feldman, 1995).

For the most part, explanations of morphological priming fall into one of three categories: *Decompositional* accounts (e.g., Taft and Forster, 1975), which assume that words are parsed into their morphological constituents in order to access the lexicon; *Whole-word* accounts (e.g., Feldman and Fowler, 1987; Lukatela *et al.*, 1980), which hold that although each word has its own lexical entry, the lexicon is organized so that morphological relationships are explicitly represented (for example, by direct connections among the entries for morphologically related words); and *Dual-process* accounts (e.g., Baayen *et al.*, 1997; Caramazza *et al.*, 1988; Stanners *et al.*, 1979), which assume that both decompositional and whole-word processes are at work. Although these accounts differ from one another in many respects, they all assume that long-term priming is a consequence of a change in the activation of a pre-existing representation, and that priming transfers between morphological relatives because the lexicon is organized around morphological principles. More generally, because they are formulated within the activation tradition described in the Introduction, all of these accounts are meant to explain not just morphological priming, but rather a variety of phenomena involving the effects of morphological structure on word identification.

The connectionist framework provides a different perspective on morphological priming—one that views morphological effects not as the consequence of the structural properties of the lexicon, but instead as the influence of statistical regularities on the dynamics of the network that is responsible for visual word identification. On this view, morphological effects stem from the fact that, with the exception of morphologically related words, similarity in word-form bears no relationship to similarity in word meaning. Hence, morphological relationships are virtually the only source of statistical regularities in the mappings between (orthographic and phonological) form and meaning. Through the covariant learning process, these regularities structure a network's weight matrices, and as a result are reflected in that network's behavior, not only in long-term morphological priming, but also in phenomena such as short-term priming (e.g., Gonnerman *et al.*, 1995), morpheme-frequency effects (e.g., Taft, 1979), and family-size effects (Schreuder and Baayen, 1997).

Put another way, from a connectionist perspective morphological effects are fundamentally interference effects: Because morphological relatives are similar in form and meaning, changing the weights to strengthen the attractor for one word also strengthens the attractors for morphologically related words. Thus, morphological effects are a manifestation of the same similarity principle that is at work in a variety of superficially unrelated phenomena. For example, just as morphological effects arise from the tendency to map similar orthographic patterns to similar semantic patterns, phonological consistency effects arise from the tendency to map similar orthographic patterns to similar phonological outputs.

Note, however, that whereas the mapping from orthography to phonology is highly structured, the mapping from form to meaning is largely arbitrary, and thus the structure imparted by morphological regularities is especially influential.

If morphological priming is a transfer effect, the degree to which priming one word facilitates a morphological relative should vary with their similarity. Thus, for example, root forms (e.g., *teach*) would be expected to be primed more strongly by regular inflections (e.g., *teaching*), which are highly similar to their root forms, than by irregular inflections (e.g., *taught*), which are usually formed by changing the orthographic and/or phonological properties of their base forms. This prediction is, in fact, consistent with experimental results (Feldman, 1994; Stanners *et al.*, 1979). However, this pattern of results can also be accounted for by models that assume that morphological relationships are explicitly represented in the lexicon, provided that different classes of morphological relationships are represented in different ways (i.e., dual-process models such as Stanners *et al.*, 1979). Thus, the contrast between priming effects involving regular and irregular inflections is not particularly diagnostic.

Thus, a more telling case involves variability in similarity when the morphological relationship between the prime and target is held constant. For example, some irregular past-tense forms are fairly similar to their root forms (e.g., *made–make, swam–swim*), whereas others are less similar (e.g., *took–take, bought–buy*). From a connectionist perspective, more morphological priming would be expected in the former case. In contrast, models that link morphological priming to the explicit representation of morphology in the lexicon would be hard-pressed to explain an effect of formal similarity within morphological class.

It turns out that formal similarity does modulate morphological priming, although the effect is sometimes rather small. Thus, experiments by Fowler *et al.* (1985), Napps (1989), Stanners *et al.* (1979), and Stolz and Feldman (1995), all failed to find an effect of orthographic similarity on morphological priming. However, in each case the numerical trend was towards less priming with less orthographic similarity, and a meta-analysis indicates that consistent trend is itself statistically reliable (Rueckl *et al.*, 1997). Moreover, all of these experiments investigated morphological priming using the lexical decision task, which may be relatively insensitive to orthographic factors. In two experiments using word fragment completion, Rueckl *et al.* (1997) found strong and statistically significant effects of orthographic similarity on the magnitude of morphological priming. Similar results have been reported by Gonnerman *et al.* (1995) concerning morphological effects in short-term priming. (Although the connectionist account assumes that different mechanisms underlie short- and long-term, the rationale for why priming should vary with similarity is essentially the same in both cases—see Gonnerman *et al.*, 1995; Plaut and Gonnerman, 2000; Raveh and Rueckl, 2000; Rueckl and Raveh, 1999.)

Another prediction of the connectionist account is that if morphological effects are due to the structure that morphological regularities impart on the mappings from form to meaning, other regularities that structure this mapping should give rise to similar sorts of effects. To test this prediction, Rueckl and Dror (1994) asked subjects to study a set of pseudoword-definition pairings over a five-week period. For some of the subjects, the pairings were constructed so that pseudowords with word bodies were systematically paired with 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 advantage of the systematic pairings reflects the manner in which these pairings structure the mapping between form and meaning. As the memory traces for learning events involving these pairings are superimposed on the network's connections, their shared structure (or lack thereof) shapes the pattern of connectivity, which in turn shapes the network's flow of activation and the behavior that manifests this flow.

Perceptual specificity effects

The previous sections were concerned with the degree to which priming one word influences the subsequent processing of other words on the basis of formal, semantic, or morphological similarity. In this section we consider transfer not from one word to another, but instead from one token of a word to perceptually distinct tokens of that same word. The question of interest is whether changing the perceptual characteristics of a word between study and test diminishes the magnitude of repetition priming. This question goes to the heart of a major theoretical divide concerning the nature of perception, memory, and cognition more generally.

On one side of the divide, developed largely within the activation tradition discussed in the Introduction, are abstractionist theories (e.g., Bowers, 1996; Grainger and Jacobs, 1996; McClelland and Rumelhart, 1981; Morton, 1979; Paap *et al.*, 1982). These theories assume that because the visual (or acoustic) details that distinguish different tokens of a word are irrelevant to that word's identity, information about these details can and should be discarded relatively early in the identification process. Thus, according to abstractionist theories, the processing of a written word results in a representation of its abstract orthographic structure in a form that is invariant over differences in case, font, and so on. Because these representations do not preserve information about visual detail, abstractionist theories predict that equivalent levels of priming should be found regardless of whether the prime and target are visually identical.

On the other side of the divide, developed largely within the memory tradition, are instance-based theories of perception and cognition (e.g., Goldinger, 1998; Jacoby and Brooks, 1984; Kolers, 1979; Whittlesea, 1987). In the case of visual word identification, these theories deny the involvement of abstract lexical representations, and hold instead that a word is identified through the retrieval of memory traces for previous processing episodes. Each trace contains a record of the processes conducted during the corresponding episode, and the contribution of each to the identification process depends on the extent to which the stimulus and other contextual cues match the information stored in that trace. Repetition priming reflects the relative accessibility of memory traces for recent events. Because these traces are thought to be highly detailed, instance theories predict that priming should be reduced by changes in perceptual detail between study and test.

A large body of experimental evidence can be marshaled in support of each view. On the one hand, the abstractionist position is supported by a number of studies which have found that a change in case (Feustel *et al.*, 1983; Scarborough *et al.*, 1977), typeface (Carr *et al.*, 1989; Rajaram and Roediger, 1993), or script (Bowers and Michita, 1998; Brown *et al.*, 1984; Feldman and Moskovljevic, 1987) has little or no effect on the size of the priming effect. The abstractionist position is also supported by an array of findings from other experimental paradigms (e.g., eye tracking, McConkie and Zola, 1979; Rayner *et al.*, 1980; masked priming, Evett and Humphreys, 1981; and CaSe MiXiNg, Besner and Johnston, 1989; Coltheart and Freeman, 1981) which suggest that word identification is driven primarily by preliminary letter identification, and not by the extraction of more holistic features that would not be invariant over differences in case or font.

Although the abstractionist position is supported by a variety of experimental results, it is not wholly consistent with the empirical data. For example, even though a number of studies have found that a change in the visual-form a word has no effect on repetition priming, a comparable number of studies have found that priming is reduced by a change in form (e.g., Brown and Carr, 1993; Jacoby and Hayman, 1987; Kolers, 1975; see Tenpenny, 1995, for a review). Similarly, although the abstractionist position is supported by the finding that the word superiority effect survives case mixing, the complex interaction of case mixing and variables such as lexicality, word frequency, and task (e.g., lexical decision, naming) has led some theorists to argue that word recognition depends, in part, on representations that preserve at least some information about visual-form, and thus are not fully invariant over differences in case and font (Allen *et al.*, 1995; Mayall *et al.*, 1997). More generally, from the perspective of a language user, information about visual detail obviously makes a difference. English, for instance, is not unusual in having a complex set of rules specifying when a word must be capitalized. In addition, writers often make use of a set of conventions that allow them to vary the look of a word in ways that carry shades of meaning—e.g., stressing a word by either italicizing it or typing it in uppercase (e.g., 'I've got a BIG problem'.). Such rules and conventions would be pointless if readers paid no attention to letter case or type font.

In sum, neither the abstractionist position nor the instance approach is wholly consistent with key experimental findings, yet each position enjoys a substantial amount of empirical (and theoretical) support. One way to resolve this apparent paradox is to adopt a two-process model that includes both abstractionist and non-abstractionist components. For example, in the memory literature, the mixed effect of perceptual format manipulations on repetition priming has been taken to indicate that two kinds of representations underlie priming, one of which is perceptually abstract and the other of which preserves information about perceptual detail (Bowers, 1996; Brown and Carr, 1993; Marsolek *et al.*, 1992). Similarly, in the word identification literature, recent accounts of the effects of case mixing have posited that preliminary letter identification is supplemented by another process that uses word-specific visual patterns to access the lexical code associated with a visual input (McClelland, 1977; Mayall *et al.*, 1997) or to assess the familiarity of that input (Besner *et al.*, 1984; Besner and Johnston, 1989).

The connectionist perspective offers a different resolution of this empirical paradox. In an important sense, the dynamics of a connectionist network gives rise to abstraction while simultaneously allowing for the preservation of visual detail (McClelland and Rumelhart, 1985). From this perspective, findings that suggest abstraction and findings that suggest preservation of detail do not reflect the operation of two complementary processes, but rather are different manifestations of the same underlying process—the activation dynamics of a network employing distributed representation, superimpositional memory, and incremental learning.

To flesh out this idea, it should first be noted that although the weight changes that result from an experience with a word are distributed throughout the lexical network, the connections of most relevance to the issue at hand are those that project to or from the so-called 'orthographic' layer. In most extant connectionist models, the representations schemes employed at this level (e.g., 'wickelgraphs', Seidenberg and McClelland, 1989; position-specific letter units, Plaut *et al.*, 1996) are abstractionist in the sense that they carry information about the identity and order of the letters that comprise a word, but not about the case, font, or other properties of those letters. Indeed, the very labeling of this layer as the 'orthographic' layer reflects the abstractionist bent of these representational schemes. Thus, for reasons that will become apparent, throughout the rest of this section I will refer to this layer as the *visual word-form*, or *VWF*, layer.

For the most part, the use of abstractionist schemes has been pragmatically motivated. Because the VWF layer typically serves as the 'input' layer, the primary constraint on the representations at this level has been that they must be appropriate for capturing the sorts of statistical regularities that characterize, for example, the mapping from orthography to phonology. This is not an unreasonable strategy, particularly given that the debate over abstraction is largely tied to findings that are outside of the scope of phenomena extant models have been intended to address. Nonetheless, the sort of representational schemes typically used fail to take advantage of one of the most interesting properties of connectionist networks—the ability of a network to develop its own representational schemes as it learns about the structure of its environment. As many modelers can attest, a network left to its own devices will usually develop a representational scheme that is an appropriate (and often surprisingly elegant) solution for the computational problem that confronts it.

Thus, to address the abstraction issue from a connectionist perspective, one should apply the maxim 'They're all hidden units'. This means that rather than treating the VWF layer as the lexical network's input layer, the VWF layer is better thought of as a hidden layer that receives connections from an input layer representing the visual properties of a stimulus (say, in terms of coarsely coded retinotopic features) and sends connections to output layers responsible for representing a word's phonological and semantic properties. If the VWF layer is treated in this way, the goal of the modeler is not to specify the VWF representations a priori, but instead to determine what sort of organization will arise given the statistical structure of the tasks that the lexical network must perform.

One important constraint on the organization of a hidden layer's representations is the similarity principle—the tendency for similar states to have similar consequences. The similarity principle implies that the organization of a hidden layer's representations will tend to reflect both the organization of the patterns of activation that evoke them (such that if two input patterns are similar, they tend to evoke similar hidden patterns) and the organization of the patterns of activation that they in turn evoke (hence, similar output patterns also tend to be associated with similar hidden patterns). (See Plaut *et al.*, 1996; Rueckl and Raveh, 1999, for analyses illustrating this point.) In the case of the VWF layer, this implies that the organization of the VWF patterns will reflect a balance of both visual (bottom-up) and phonological and semantic (top-down) influences. The bottom-up constraints pressure the network to map similar visual inputs to similar VWF patterns. Conversely, the top-down constraints pressure the network to assign similar hidden patterns to input patterns that are similar in *function* (i.e., are mapped to similar phonological and semantic output representations). In some cases the bottom-up and top-down forces act in concert—that is, some visually similar forms (e.g., R/R, C/c) are generally mapped to the same phonological and semantic outputs. However, in many cases the bottom-up and top-down constraints are at odds: Visually similar forms (e.g., r/n) may be functionally dissimilar, and functionally similar forms (e.g., R/r) may be visually dissimilar.

Note that in some ways these constraints tend to that favor the adoption of an abstractionist representational scheme, but that in other ways they exert pressure in favor of the preservation of perceptual detail. With the exception of the relatively rare cases where differences in visual-form carry information about meaning (e.g., *Mark* versus *mark*, *Penny* versus *penny*), top-down constraints generally favor abstraction in the VWF layer. That is, these constraints pressure the system to assign similar VWF patterns to visual inputs that have similar phonological and semantic properties, even if they are visually dissimilar. In contrast, because bottom-up constraints pressure the system to map visually similar inputs to similar VWF patterns, these constraints often work against the tendency for abstraction by bringing together the hidden representation for functionally distinct inputs (e.g., *r* versus *n*) and pulling apart the hidden representations for functionally equivalent inputs (e.g., *R* versus *r*). Thus, the similarity principle implies that the organization of the VWF representations will be influenced by both constraints that pressure the system towards abstraction and constraints that pressure it towards an organization that preserves perceptual detail, and thus that the degree to which a network's VWF representations are perceptually abstract or perceptually detailed depends on the balance between these constraints.

A related principle provides further clarification of this issue. According to the principle of *quasi-equivalent states*, distinct patterns of activation can have the same behavioral consequences. In fact, due to both the linear transformations that occur when a pattern of activation is projected over a set of weights and the non-linear transformations that occur when the resulting excitatory and inhibitory signals are passed through an activation function, in some cases distinct patterns of activation at one layer result in responses that are literally indistinguishable at the next. More generally, however, if two patterns of activation at one level evoke highly similar responses, these patterns may be for all intents and purposes indistinguishable to an outside observer (hence the term '*quasi-equivalent*').

The principle of quasi-equivalent states implies that although the top-down constraints favoring abstraction must be honored, they need not completely override the bottom-up constraints favoring an organization based on visual, rather than functional, similarity. Consequently, although it is appropriate to think about the representation of a particular *token* of a word as a pattern of activation, or point in the network's state space, the representation of a word *type* is better thought of as a region of such points—a 'region of functional equivalence' (see Elman, 1995). Visually distinct tokens of the same word (e.g., BAR, bar) can be represented by different points within this region, and yet the network can respond appropriately to each. If the two patterns are similar enough, the behaviors they drive will be virtually identical. As their differences grow, however, they may well have different behavioral consequences. Such consequences might occur, for example, if they differ in the efficiency with which they produce their associated phonological and semantic output patterns, even though, by assumption, they ultimately produce identical output patterns.[8] On this view, abstraction is not all-or-none but a matter of degree. Whether the representations involved in word identification appear to be abstract or perceptually detailed depends on both the similarity of the VWF patterns activated by different visual-forms and the measure of performance used to assess a reader's internal state.

This perspective on the representation of visual word-form suggests a way to reconcile the conflicting patterns of results described above. Advocates of the abstractionist and instance approaches have often attempted to reconcile these patterns by downplaying the implications of one set of results or the other. For example, proponents of abstractionist accounts (e.g., Bowers, 1996; Carr *et al.*, 1989; Ratcliff and McKoon, 1997) have sometimes argued that specificity effects usually occur under relatively atypical circumstances (e.g., when stimuli are presented in highly unusual fonts), and that these effects therefore reflect the operation of a supplementary process that plays little role in 'normal' reading. Conversely, proponents of episodic theories of priming (e.g., Jacoby and Hayman, 1987; Tenpenny, 1995) have downplayed evidence favoring the abstractionist position by pointing out that this evidence rests on the acceptance of a null effect and that, in most cases where the specificity effect is not statistically significant, the trend is in the right direction.

From the connectionist perspective, the claim that specificity effects have little to do with 'normal' reading, like the claim that pseudoword identification has little to do with word identification, rings hollow. At the same time, although it is certainly true that null effects must be treated cautiously, the fact that specificity effects often fail to occur must be given some weight. Thus, it seems likely that both patterns of results hold important clues about the nature of the processes that underlie word perception. Indeed, it is no accident that the word identification system seems to operate in an abstractionist mode in some circumstances and a non-abstractionist mode in others—these patterns of behavior reflect the fact that the dynamics of visual word identification are under the influence of a multiplicity of factors.

Consider, for example, the effect of script familiarity on specificity effects in priming. As several authors have pointed out (Bowers, 1996; Brown and Carr, 1993; Tenpenny, 1995), specificity effects seem to occur more reliably when the target stimuli have relatively novel surface forms. Specificity effects are often found when the target words are handwritten (Brown and Carr, 1993), printed in an unusual font (e.g., Graf and Ryan, 1990; Jacoby and Hayman, 1987), or presented in alternating case (Brooks, 1978). In contrast, when the target words are printed in a standard type font, specificity effects often fail to occur (e.g., Brown and Carr, 1993; Levy and Kirsner, 1989; Scarborough *et al.*, 1977), even if the primes are presented in relatively novel forms. Thus, there is an asymmetry in the transfer of

---

[8] With a sufficiently large distance between the patterns, of course, the patterns will fall into different regions of equivalence, and hence will be mapped onto different output patterns.

priming between typical and atypical visual-forms: Relative to the same-form conditi atypical forms prime typical forms more fully than the reverse.

This asymmetry in the transfer of priming between typical and atypical visual-forms be understood in terms of the effect of learning on the dynamics of a system that is bet attuned for the processing of typical forms. Because a stimulus printed in a standard t font is more representative of the experiences that have structured the network's weigh word written in a typical script will be represented by an attractor near the center of word's region of functional equivalence, and consequently, because it benefits from a r tively large number of past experiences, the activation dynamics stabilize relatively quic In contrast, a word written in an unusual form will be represented by an attractor on periphery of its region of functional equivalence, and because it benefits less from experiences, the activation dynamics take longer to settle. As a result of these difference the 'baseline' dynamics, the effect of a single additional learning event will have a grea impact on the processing of atypical forms than on the processing of typical forms, anc well, the effect of differences in the visual-form of the prime and target will be magnifi Thus, specificity effects should be more readily observed when the target is presented in unusual form, even though (in an absolute sense) a change in the form between study a test must always result in a reduction in priming.

It is worth noting that each facet of this account is consistent with extant data. Wo printed in unusual forms are generally more difficult to process (Adams, 1979; Allen *et* 1995; Besner *et al.*, 1984; Brown and Carr, 1993; McClelland, 1976). As well, in experime where a specificity effect is obtained in one condition but not another, the magnitude same-form priming is usually greater in the condition where the specificity effect v found (e.g., Bowers, 1996; Brown and Carr, 1993; Marsolek *et al.*, 1992). Finally, specific effects are sometimes found with stimuli presented in standard fonts (Blaxton, 19 Jacoby and Hayman, 1987), and as noted above, when the specificity effect does not rea statistical significance, the numerical difference between the same- and different-fo conditions is consistently in the expected direction (Tenpenny, 1995). This pattern is cc sistent with the position that the effect of script familiarity on the specificity effect quantitative rather than qualitative.

Although script typicality appears to play an especially critical role in determining magnitude of specificity effects, other factors may also be relevant. For example, specific effects appear to be larger for low-frequency words than for high-frequency words (Jacc and Hayman, 1987), and for pseudoword targets than for word targets (Bowers, 19 Brown and Carr, 1993).[9] Because more encounters with an input tend to increase 'center of mass' of a word's region of functional equivalence, attractors for tokens of a hi frequency words will tend to gravitate toward the center of that word's region more so th will the attractors for low-frequency words or pseudowords. Thus, high-freque

---

[9] Brown and Carr (1993) concluded that specificity effects were equivalent for words and pseudowords. Howe their conclusion was based on analyses combining lexical decision and speeded naming data, and as noted in earlier section, the lexical decision task may be an inappropriate task for studying pseudoword priming. In f in the lexical decision task, Brown and Carr did not obtain a significant pseudoword priming effect. When o the naming results are considered, the specificity effect was about 11 ms for pseudowords and about 2 ms words.

words will tend to be represented by relatively 'abstract' representations, whereas the representations of pseudowords will tend to preserve perceptual detail. Moreover, because lexicality and word frequency influence an item's baseline settling time, they modulate the sensitivity to specificity manipulations in the same manner as does script typicality.

Manipulations of task demands have also been associated with differences in specificity effects in repetition priming. For example, with other factors held constant, a change in visual format reduced priming when the study task emphasized perceptual processing, but not when the study task focused on semantic properties (Blaxton, 1989; Curran *et al.*, 1996; Graf and Ryan, 1990). It is not hard to see how task demands might affect the dynamics of word identification to give rise to this pattern of results. Network models often include 'gain' parameters that specify the slope of the activation function computed by nodes within a layer or the rate with which this activation is transmitted to other layers (Farrar and Van Orden, submitted; Hinton and Sejnowski, 1986; Plaut *et al.*, 1996). These parameters could easily be tuned in accordance with task demands. For example, when the task draws attention to the perceptual characteristics of the stimulus, the control parameters might be set so that the bottom-up influence on the dynamics is relatively strong. Conversely, if the task emphasizes higher-level processes, the gain parameters at these levels might be increased, shifting the balance towards more top-down influence on the activation dynamics. As the balance between the bottom-up and top-down influences on the activation dynamics shifts with task demands, so too would the likelihood that priming would be affected by a change in visual-form.

A final factor that should be considered with regard to specificity effects is hemispheric specialization. In a series of studies, Marsolek and others (Marsolek *et al.*, 1992, 1994; Marsolek and Burgund, 1997; Marsolek and Hudson, 1999) have found that specificity effects are larger when the right hemisphere plays a relatively large role in visual processing (as a consequence of lateralized presentation, for example). Because network models of letter and word processing have rarely taken cerebral lateralization into account (see Shevtsova and Reggia, 1999, for an exception), these findings pose a challenge for the connectionist approach. One possibility is to modify the structure of the triangle model so that the phonological and semantic pathways receive input from two layers of nodes corresponding to visual word-form areas in the left and right hemisphere. If, as the data suggest, the VWF area in the left hemisphere tends to adopt a more abstractionist organization than the corresponding right hemisphere structure, one could ask whether this difference should be attributed to an intrinsic hemispheric difference in, say, activation functions or initial patterns of connectivity, or whether instead it is a consequence of hemispheric differences in the nature of the bottom-up input or top-down feedback received by these areas.

It is worth noting that if a network model posits two distinct visual word-form areas to account for the hemispheric differences in perceptual specificity effects, it would in some sense be a 'two-process' model, but not of the sort favored by proponents of the abstractionist position (e.g., Bowers, 1996; Brown and Carr, 1993). In particular, in the network model it would not be the case that, strictly speaking, the left hemisphere uses abstract representations and the right maintains perceptual detail. Instead, given the constraints that determine a network's activation dynamics, in both hemispheres abstraction would necessarily be a matter of degree. Thus, even the left hemisphere could give rise to

specificity effects under some circumstances, as suggested by the results of Marsolek an Hudson (1999).

In summary, from the connectionist perspective it is no accident that experimenta results sometimes suggest that readers make use of abstract representations and sometime suggest that perceptual detail matters. To a certain extent, these contrasting patterns o results reflect the limitations of our methodologies—how much detail we can see in th representations underlying visual word identification depends in part on the resolvin power of the instruments we use to observe them. More importantly, however, because o the conflicting constraints imposed by the similarity principle, these representations ar organized such that different tokens of a word are mapped onto similar, but generally no identical, patterns of activation. Thus, abstraction is neither all nor none, but instead is matter of degree.

Summary: network models and other accounts

The purpose of this chapter is to provide an overview of the connectionist perspective or repetition priming in visual word identification. From this perspective word identification like other cognitive tasks, is accomplished by a network of simple processing units tha interact by sending each other inhibitory and excitatory signals. A learning process shape these interactions by adjusting the strengths of the connections among the nodes, thu attuning the behavior of the network to the regularities and task demands of its environment. Repetition priming is a manifestation of this learning process, revealing how th dynamical behavior of the network changes as a consequence of a single learning event.

The connectionist framework incorporates several major assumptions about the structure and operation of the networks underlying cognitive tasks. One of these is tha representations are distributed across many nodes, and thus that each node plays a role i many different representations. Inherent in the notion of distributed representation is th fact that some patterns are more similar than others. Thus, for example, orthographicall similar words are represented by relatively similar patterns of activation at the visual word-form level; the representations at this level are even more similar for morphologicall related words, and more similar still for different tokens of the same word.

One of the major themes running throughout this chapter is that the core assumptions of the connectionist framework leave their fingerprints on the characteristics of repetition priming. For example, because distributed representation is coupled with superimpositional storage, repetition priming gives rise to transfer effects—modifying the weights to improve the response to one word changes the manner in which the network responds to other words as well. Similarly, because the similarity principle and the principle of quasi-equivalent states imply that different tokens of a word are represented by different but highly similar patterns of activation, perceptual abstraction is a matter of degree. As a result, perceptual specificity effects come and go depending on various aspects of the experimental circumstance.

Another major theme running throughout this chapter is that there is a deep and intimate relationship between repetition priming and other classes of behavioral phenomena. Repetition priming takes its form due to the same mechanisms and principles that shape other aspects of behavior. Thus, both transfer effects in priming and consistency effects in naming can be understood as manifestations of the similarity principle; morphological effects in long-term priming and morphological effects in short-term priming both reflect the manner in which statistical regularities in the environment structure a network's weights, and hence its activation dynamics; neither the processes responsible for identification nor the processes responsible for priming distinguish between words and pseudowords.

In this light, it is worth considering the relationships between the connectionist approach and the two theoretical traditions discussed at the beginning of the chapter. Like the memory tradition, the connectionist approach views repetition priming as a form of learning, and explains priming in terms of principles that apply to a wide variety of phenomena. Like the activation tradition, the connectionist approach provides a mechanistic account of repetition priming, and places this account in the context of an explanatory framework that explicitly addresses other classes of phenomena related to specific perceptual tasks. Thus, to a certain extent the connectionist approach represents a blend of the activation and memory traditions.

However, it is also the case that the connectionist approach makes some assertions that either conflict with other kinds of accounts or concern issues about which these accounts are relative mute. Yet, even on these issues it is possible to see potential avenues of convergence. For example, instance theories (e.g., Goldinger, 1998; Hintzman, 1986; Jacoby and Brooks, 1984; Kolers, 1979; Whittlesea, 1987) have emerged as one important class of models within the memory tradition. To the extent that these models are described in terms of explicit encoding and retrieval processes, they generally posit mechanisms that are rather unlike the activation and learning processes assumed by connectionist models. Yet, instance models and connectionist models share some common ground—both deny the psychological reality of abstract representations (at least as they are commonly understood), both emphasize the notion of 'representations' as emergent states, and both assume that every experience can potentially play a role in the emergence of these states. An exploration of these points of agreement may reveal that the primary insights of the instance approach can be understood in terms of the dynamics of a connectionist network, and conversely, these insights may provide the basis for a deeper understanding of these dynamics.

A similar point can be made with regard to some of the structural models that have emerged from the memory tradition. In particular, multiple-memory-systems accounts (e.g., Marsolek *et al.*, 1992; Schacter, 1992; Squire, 1992) hold that the dissociations between priming and other forms of memory (as well as dissociations among forms of priming) occur because different neural subsystems are responsible for different kinds of processes. Typically, systems accounts focus on the macrolevel: They are more concerned with determining the functional partitioning of the brain and identifying the role of each subsystem than with providing an explicit account of how each subsystem accomplishes its task. Connectionist models can complement the systems approach by offering insights about how each subsystem operates, thus clarifying, for example, why certain functions appear to be 'computationally incompatible' (Rueckl *et al.*, 1989; Marsolek *et al.*, 1992; Marsolek and Burgund, 1997). Conversely, to a certain extent there is little justification for the architectural assumptions of current connectionist models. For example, the architecture of the triangle model (Harm, 1998; Plaut *et al.*, 1996; Seidenberg and McCelland, 1989) seems to be based more on intuition than on solid theoretical or empirical grounds, and in fact other kinds of architectures have been explored with some success (Kello and Plaut, submitted; Zorzi *et al.*, 1998). Thus, theorists working within the systems approach may be able to provide connectionist modelers with useful theoretical insights and empirical constraints. (For an example, see the discussion of hemispheric specialization in the section titled *Perceptual specificity effects.*)

Finally, while the connectionist approach is clearly more at odds with activation theories than with theories developed within the memory tradition, there are ways in which the gap between the two kinds of approaches might be bridged. One of the main conflicts between these approaches lies in their conceptions of mental representations—as either localist, abstract entities or distributed patterns of activation. Perhaps it would be possible to think about the representations in an activation model as shorthand descriptions for the attractors in a connectionist network. One benefit of this approach is that the flow of activation within a localist network is relatively easy to track and interpret. Thus, the connectionist approach might gain in its understanding of the cooperative and competitive forces that underlie cognition by relating the dynamics of a distributed system to, say, the excitatory and inhibitory interactions among word nodes (Grainger and Jacobs, 1996; McClelland and Rumelhart, 1981).

Another major conflict between the connectionist and activation approaches is more meta-theoretical in nature. From the connectionist perspective, one of the least appealing aspects of the activation approach is its tendency to mirror the complexity of the behavior it intends to explain in the complexity of the models that it offers as explanations. It is an exaggeration to claim that abstractionist theories are constructed in accordance with the maxim 'For each effect posit a new cause', but it is less of an exaggeration than one might like. According to models developed within the abstractionist position: Different processes underlie word and pseudoword priming; different processes are needed to read *pint* and *mave* aloud; morphologically complex words can be read in either of two ways; morphologically complex words can be produced in either of two ways; priming is sometimes due to representations that are perceptually abstract, and other times due to representations that preserve perceptual detail. And so on.

The point here is not so much about the sheer complexity of the organization of the cognitive processes posited by activation accounts, although arguably the reliance on symbolic representations and symbol manipulating processes introduces more complexity than is actually needed. The point instead is that activation accounts rarely attempt to explain where this organization comes from. If each word in a reader's vocabulary is represented by a logogen, how did these logogens get into the reader's head in the first place? If the architecture of the reading system includes two different routes that make use of fundamentally different kinds of processes, how did this architecture get put into place? In the words of the philosopher Dan Dennett, a theory 'takes out a loan on intelligence' when it takes the organization of a cognitive processes as a given. A loan on intelligence is fine if it is paid back with an explanation of how that organization came about. The worry about the activation tradition is that generally it has exhibited little if any concern with paying back its loans. It has the flavor of a Reaganomics approach to cognitive theories.

In contrast, connectionism's emphasis on learning and self-organization is an acknowledgement that the need to explain the organization of psychological processes is as important as the need to explain the organization of the behaviors that those processes generate. In the connectionist approach richly organized activation dynamics are a consequence of both the interactions of processing units that follow a simple algebraic rule and the pattern of connectivity that constrains these interactions. The pattern of connectivity is itself the product of a dynamic process (at a slower time scale) involving a simple algebraic rule and the constraints provided by the environment inhabited by the network. A major theme of this chapter (and of the relevant studies cited within) has been to demonstrate how these self-organizing dynamics give rise to the effects of repetition priming, phonological consistency, morphological complexity, lexicality, word frequency, and so forth.

To be clear: The claim is not that connectionist models fully explain the organization of cognitive processes and thus don't take out loans on intelligence. Rather, the claim is that a major emphasis of the connectionist approach is to minimize the need for such loans and to pay back the ones that are taken out. For example, simulations of the triangle model of word identification (Plaut *et al.*, 1996; Seidenberg and McClelland, 1989) make assumptions about the input and output representations, the teacher (i.e., the source of the target patterns that drive error-correction learning), and the structure of the architecture. Paying back these loans forms part of the connectionist research agenda, and it is worth noting that possible means of payback for each assumption can be identified:

(a) In the long run, the input and output units of the triangle model can (and must) be treated as hidden units, and thus the principles that apply to the organization of hidden unit representations will apply here as well (see the section on *Perceptual specificity effects*).

(b) In some cases a network can use its own behavior as the basis for an error signal (O'Reilly, 1996); alternatively, some algorithms do not require an error signal to drive learning (Grossberg, 1987; Stark and McClelland, 2000).

(c) In some learning algorithms regions of an initially undifferentiated network become specialized for specific computational tasks, such that a more-or-less modular architecture emerges with experience (e.g., Jacobs and Jordan, 1992; see Elman *et al.*, 1996, for a broad discussion).

Of course, there is no guarantee that any of these possibilities will ultimately prove fruitful. Only time and hard work will tell. Even so, these issues and avenues of research illustrate two of the central—and to me, most attractive—aspects of the connectionist approach: The desire for theories that explain, rather than assume, the organization of mental processes, and a rich set of theoretical constructs that are conducive to the development of explanations of this sort.

Acknowledgments

The preparation of this chapter, and some of the research described herein, were supported by National Institute for Child Health and Development grant HD-01994 to Haskins Laboratories.

References

Adams, M. J. (1979). Models of word recognition. *Cognitive Psychology*, **11**, 133–76.

Allen, P. A., Wallace, B., and Weber, T. A. (1995). Influence of case type, word frequency, and exposure duration on visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 914–34.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin and Review*, **4**, 439–61.

Baayen, R. H., Dijkstra, T., and Schreuder, R., (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, **37**, 94–117.

Becker, S., Moscovitch, M., Behrmann, M., and Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23**, 1059–82.

Bentin, S. and Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology*, **42A** (4), 693–711.

Besner, D. and Johnston, J. C. (1989). Reading and the mental lexicon: On the uptake of visual information. In *Lexical processes and representation* (ed. W. Marslen-Wilson), pp. 291–316. Cambridge, MA: MIT Press.

Besner, D., Davelaar, E., Alcott, D., and Perry, P. (1984). Wholistic reading of alphabetic print: Evidence from the FBM and the FBI. In *Orthographies and reading* (ed. L. Henderson), pp. 121–35. Hillsdale, NJ: Erlbaum.

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 657–88.

Booker, J. (1992). *Interference effects in implicit and explicit memory*. Unpublished doctoral dissertation. University of Arizona.

Bowers, J. S. (1994). Does implicit memory extend to legal and illegal nonwords? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 534–49.

Bowers, J. S. (1996). Different perceptual codes support priming for words and pseudowords: Was Morton right all along? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **22**, 1336–53.

Bowers, J. S. (1999). Priming is not all bias: Commentary on Ratcliff and McKoon (1997). *Psychological Review*, **106**, 582–96.

Bowers, J. S. and Michita, Y. (1998). An investigation into the structure and acquisition of orthographic knowledge: Evidence from cross-script Kanji-Hiragana priming. *Psychonomic Bulletin and Review*, **5**, 259–64.

Bowers, J. S., Damian, M., and Havelka, J. (in press). Can distributed orthographic knowledge support word specific long-term priming? Apparently so. *Journal of Memory and Language*.

Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In *Cognition and categorization* (ed. E. Rosch and B. B. Lloyd), pp. 169–211. New York: Wiley.

Brown, H. L., Sharma, N. K., and Kirsner, K. (1984). The role of script and phonology in lexical representation. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **36A**, 491–505.

Brown, J. S. and **Carr, T. H.** (1993). Limits on perceptual abstraction in reading: Asymmetric transfer between surface forms differing in typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **19**, 1277–96.

Burani, C. and **Caramazza, A.** (1987). Representation and processing of derived words. *Language and Cognitive Processes*, 2, 217–27.

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

Carr, T. H., Brown, J. S., and **Charalambous, A.** (1989). Repetition and reading: Perceptual encoding mechanisms are very abstract but not very interactive. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 763–78.

Cermak, L. S., Talbot, N., Chandler, K., and **Wolbarst, L.** (1985). The perceptual priming phenomenon in amnesia. *Neuropsychologia*, **23**, 615–22.

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

Coltheart, M. and **Freeman, N.** (1974). Case alternation impairs word identification. *Bulletin of the Psychonomic Society*, 3, 102–4.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and **Ziegler, J.** (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, **108**, 204–56.

Cree, G. S., McRae, K., and **McNorgan, C.** (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, **23**, 371–414.

Curran, T., Schacter, D. L., and **Bessenoff, G.** (1996). Visual specificity effects on word stem completion: Beyond transfer appropriate processing? *Canadian Journal of Psychology*, **50**, 22–33.

Danenbring, G. L. and **Briand, K.** (1982). Semantic priming and the word repetition effect in a lexical decision task. *Canadian Journal of Psychology*, **36**, 435–44.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**, 17–22.

Diamond, R. and **Rozin, P.** (1984). Activation of existing memories in the amnesic syndrome. *Journal of Abnormal Psychology*, **93**, 98–105.

Dorfman, J. (1994). Sublexical components in implicit memory for novel words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 1108–25.

Elman, J. L. (1995). Language as a dynamical system. In *Mind as motion: Explorations in the dynamics of cognition* (ed. R. F. Port and T. van Gelder), pp. 195–226. Cambridge, MA: MIT Press.

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

Evett, L. J. and **Humphreys, G. W.** (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology*, **33A**, 325–50.

Farrar, W. T. and **Van Orden, G. C.** (submitted). Nonlinear dynamics of naming errors. Manuscript submitted for publication.

Feldman, L. B. (1991). The contribution of morphology to word recognition. *Psychological Research*, **53**, 33–41.

Feldman, L. B. (1994). Beyond orthography and phonology: Differences between inflections and derivations. *Journal of Memory and Language*, **33**, 442–70.

Feldman, L. B. and **Fowler, C. A.** (1987). The inflected noun system in Serbo-Croatian: Lexical representation of morphological structure. *Memory and Cognition*, **15** (1), 1–12.

Feldman, L. B. and **Moskovljevic, J.** (1987). Repetition priming is not purely episodic in origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 573–81.

Feldman, L. B., Rueckl, J. R., DiLiberto, L., Pastizzo, M., and **Vellutino, F. R.** (in press). Morphological analysis by child readers as revealed by the fragment completion task. *Psychonomic Bulletin and Review.*

Feustel, T. C., Shiffrin, R. M., and **Salasoo, A.** (1983). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General*, **112**, 309–46.

Forbach, G., Stanners, R., and **Hochhaus, L.** (1974). Repetition and practice effects in a lexical decision task. *Memory and Cognition*, 2, 337–9.

Forster, K. (1987). Form-priming with masked primes: The best match hypothesis. *Attention and Performance XII*, 127–45.

Fowler, C. A., Napps, S. E., and **Feldman, L. B.** (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory and Cognition*, **13**, 241–55.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 674–91.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, **105**, 251–79.

Gonnerman, L., Devlin, J., Andersen, E. S., and **Seidenberg, M. S.** (1995). 'Morphological' priming without a morphological level of representation. *Journal of the International Neuropsychological Society*, **1**, 142.

Graf, P. and **Ryan, L.** (1990). Transfer-appropriate processing for implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 978–92.

Graf, P. and **Schacter, D. L.** (1987). Selective effects of interference on implicit and explicit memory for new associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 45–53.

Grainger, J. and **Jacobs, A. M.** (1996). Orthographic processing in visual word recognition: A multiple readout model. *Psychological Review*, **103**, 518–65.

Grainger, J. and **Jacobs, A. M.** (1998). On localist connectionism and psychological science. In *Localist connectionist approaches to human cognition* (ed. J. Grainger and A. M. Jacobs), pp. 1–38. Mahwan, NJ: Lawrence Erlbaum Associates.

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science*, **11**, 23–63.

Grossberg, S. and **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. Unpublished dissertation: Southern California University.

Harm, M. W. and **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 *Progress in the psychology of language* (ed. A. W. Ellis), Vol. 1, pp. 15–72. London: Erlbaum.

Hetherington, P. A. and **Seidenberg, M. S.** (1989). Is there 'catastrophic interference' in connectionist networks? In *Proceedings of the 11th Annual Meeting of the Cognitive Science Society*. pp. 26–33. Hillsdale, NJ: Erlbaum.

Hinton, G. E. and **Sejnowski, T. J.** (1986). Learning and relearning in Boltzmann machines. In *Parallel distributed processing: Explorations in the microstructure of cognition* (ed. D. E. Rumelhart and J. L. McClelland), Vol. 1, Foundations. Cambridge, MA: MIT Press.

Hintzman, D. L. (1986). 'Schema abstraction' in a multiple-trace memory model. *Psychological Review*, **93**, 411–28.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational properties. *Proceedings of the National Academy of Sciences USA*, **79**, 2554–8.

Humphreys, G. W., Besner, D., and **Quinlan, P. T.** (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General*, **117**, 51–67.

Jacobs, R. A. and **Jordan, M. I.** (1992). Computational consequences of a bias towards short connections. *Journal of Cognitive Neuroscience*, **4**, 323–6.

Jacoby, L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **9**, 21–38.

Jacoby, L. L. and **Brooks, L. R.** (1984). Nonanalytic cognition: Memory, perception, and concept learning. In *The psychology of learning and motivation: Advances in research and theory* (ed. G. H. Bower), Vol. 18, pp. 1–47. New York: Academic Press.

Jacoby, L. L. and **Dallas, M.** (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306–40.

Jacoby, L. L. and **Hayman, C. A. G.** (1987). Specific visual transfer in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 456–63.

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

Joordens, S. and **Becker, S.** (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23**, 1083–105.

Joordens, S. and **Besner, D.** (1994). When banking on meaning is not (yet) money in the bank: Explorations in connectionist modeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 1051–62.

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.

Kawamoto, A. H., Farrar, W. T., and **Kello, C.** (1994). When two meanings are better than one: Modeling the ambiguity advantage using a recurrent distributed network. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 1233–48.

Keane, M. M., Gabrieli, J. D., Growdon, J. H., and **Corkin, S.** (1994). Priming in perceptual identification of pseudowords is normal in Alzheimer's disease. *Neuropsychologia*, **3**, 343–56.

Keane, M. M., Verfaellie, M., Gabrieli, J. D. E., and **Wong, B. M.** (2000). Bias effects in perceptual identification: A neuropsychological investigation of the role of explicit memory. *Journal of Memory and Language*, **43**, 316–44.

Kello, C. T. and **Plaut, D. C.** (submitted). Strategic control by gain manipulation in a computational model of word reading in the tempo naming task.

Kirsner, K. and **Smith, M.** (1974). Modality effects in word identification. *Memory and Cognition*, **2**, 637–40.

Kolers, P. A. (1975). Specificity of operations in sentence recognition. *Cognitive Psychology*, **1**, 289–306.

Kolers, P. A. (1979). A pattern-analyzing basis for recognition memory. In *Levels of processing and human memory* (ed. L. S. Cermak and F. I. M. Craik). Hillsdale, NJ: Erlbaum.

Levy, B. A. and **Kirsner, K.** (1989). Reprocessing text: Indirect measures of word and message level processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 407–17.

Lukatela, G., Gligorijevic, B., Kostic, A., and **Turvey, M. T.** (1980). Representation of inflected nouns in the internal lexicon. *Memory and Cognition*, **8**, 415–23.

Lustig, C. and **Hasher, L.** (2001). Implicit memory is vulnerable to proactive interference. *Psychological Science*, **12**, 408–12.

Mandler, G., Graf, P., and **Kraft, D.** (1986). Activation and elaboration effects in recognition and word priming. *Quarterly Journal of Experimental Psychology*, **38A**, 645–62.

Mandler, G., Hamson, C. O., and **Dorfman, J.** (1990). Tests of dual process theory: Word priming and recognition. *Quarterly Journal of Experimental Psychology*, **42A**, 713–39.

Marsolek, C. J. and **Burgund E. D.** (1997). Computational analyses and hemispheric asymmetries in visual-form processing. In *Cerebral asymmetries in sensory and perceptual processing* (ed. S. Christman), pp. 125–58. Amsterdam: Elsevier.

Marsolek, C. J. and **Hudson, T. E.** (1999). Task and stimulus demands influence letter-case specific priming in the right cerebral hemisphere. *Laterality*, **4**, 127–47.

Marsolek, C. J., Kosslyn, S. M., and **Squire, L. R.** (1992). Form-specific visual priming in the right cerebral hemisphere. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **18**, 492–508.

Marsolek, C. J., Squire, L. R., Kosslyn, S. M., and **Lulenski, M.** (1994). Form-specific explicit and implicit memory in the right cerebral hemisphere. *Neuropsychology*, **8**, 588–97.

Masson, M. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21** (1), 3–23.

Masson, M. E. J. (1999). Semantic priming in a recurrent network: Comment on Dalrymple-Alford and Marmurek (1999). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25**, 776–94.

Mayall, K. A., Humphreys, G. W., and **Olson, A.** (1997). Disruption to word or letter processing? The origins of case-mixing effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23**, 1275–86.

Mayes, A., Pickering, A., and **Fairbairn, A.** (1987). Amnesic sensitivity to proactive interference: its relationship to priming and the causes of amnesia. *Neuropsychologia*, **25**, 211–20.

McClelland, J. L. (1976). Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, **4**, 80–91.

McClelland, J. L. (1977). Letter and configurational information in word identification. *Journal of Verbal Learning and Verbal Behavior*, **16**, 147–50.

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

McClelland, J. L. and **Rumelhart, D. E.** (1985). Distributed memory and the representation of general and specific knowledge. *Journal of Experimental Psychology: General*, **114**, 159–88.

McClelland, J. L., McNaughton, B. L., and **O'Reilly, R. C.** (1995). Why there are complimentary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, **102**, 419–57.

McCloskey, M. and **Cohen, N.** (1989). Catastrophic interference in connectionist networks: The sequential learning problem. *The Psychology of Learning and Motivation*, **24**, 109–65.

McConkie, G. and **Zola, D.** (1979). Is visual information integrated across successive fixations in reading? *Perception and Psychophysics*, **25**, 221–4.

McDermott, K. (1997). Priming on perceptual implicit memory tests can be achieved through presentation of associates. *Psychonomic Bulletin and Review*, **4**, 582–6.

Minsky, M. and **Papert, S.** (1969). *Perceptrons*. Cambridge, MA: MIT Press.

Monsell, S. (1985). Repetition and the lexicon. In *Progress in the psychology of language* (ed. A. W. Ellis), Vol. 2. London: Erlbaum.

Morton, J. (1969). Interaction of information in word identification. *Psychological Review*, **76**, 165–78.

Morton, J. (1979). Facilitation in word recognition: Experiments causing a change in the logogen model. In *Processing visible language I* (ed. P. A. Kolers, M. E. Wrostal, and I. Bouma), pp. 259–68. New York: Plenum Press.

Murre, J. (1992). The effects of pattern presentation on interference in backpropagation networks. In *Proceedings of the 14th Annual Meeting of the Cognitive Science Society*, pp. 54–9. Hillsdale, NJ: Erlbaum.

Murrell, G. A. and **Morton, J.** (1974). Word recognition and morpheme structure. *Journal of Experimental Psychology*, **102**, 963–8.

Napps, S. E. (1989). Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory and Cognition*, **17**, 729–39.

Napps, S. E. and **Fowler, C. A.** (1987). Formal relationships among words and the organization of the mental lexicon. *Journal of Psycholinguistic Research*, **16**, 257–72.

Neaderhiser, B. J. and **Church, B. A.** (2000, April). Priming in forced choice identification: Memory or bias? Paper presented at the 92nd Meeting of the Southern Society for Philosophy and Psychology, Atlanta.

Neisser, U. (1954). An experimental distinction between perceptual processes and verbal response. *Journal of Experimental Psychology*, **47**, 399–402.

Nelson, D., Keelean, P., and **Negrao, M.** (1989). Word-fragment cueing: The lexical search hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **15**, 333–97.

O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, **8** (5), 895–938.

Overson, C. and **Mandler, G.** (1987). Indirect word priming in connected and phonological contexts. *Bulletin of the Psychonomic Society*, **25**, 229–32.

Paap, K. R., Newsome, S. L., McDonald, J. E., and **Schvanevelt, R. W.** (1982). An activation-verification model for letter and word recognition: The word superiority effect. *Psychological Review*, **89**, 573–94.

Pinker, S. and **Prince, A.** (1988). On language and connectionism: Analysis of a parallel distributed processing model off language acquisition. *Cognition*, **28**, 73–193.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, **12**, 7[illegible]–808.

Plaut, D. C. and **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. and **Gonnerman, L. M.** (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes*, **15**, 445–85.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., and **Patterson, K.** (996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, **103**, 56–115.

Rajaram, S. and **Roediger, H. L.** (1993). Direct comparison of four implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **19** (4), 76–76.

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

Ratcliff, R. and **McKoon, G.** (1996). Bias effects in implicit memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 754–67.

Ratcliff, R. and **McKoon, G.** (1997). A counter model for implicit priming in perceptual word identification. *Psychological Review*, **104**, 319–43.

Ratcliff, R., Hockley, W., and **McKoon, G.** (1985). Components of activation: Repetition and priming effects in lexical decision and recognition. *Journal of Experimental Psychology: General*, **114**, 435–50.

Raveh, M. and **Rueckl, J.** (2000). Equivalent effects of inflected and derived primes: Long-term morphological priming in fragment completion and lexical decision. *Journal of Memory and Language*, **42**, 103–19.

Rayner, K., McConkie, G., and **Zola, D.** (1980). Integrating information across eye movements. *Cognitive Psychology*, **12**, 206–26.

Richardson-Klavehn, A. and **Bjork, R. A.** (1988). Measures of memory. *Annual Review of Psychology*, **39**, 475–543.

Roediger, H. L. (1990). Implicit memory: Retention without awareness. *American Psychologist*, **45**, 1043–56.

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

Ross, S., Yacowzer, M., and **Williams, G. M.** (1956). Recognitive thresholds for words as a function of set and similarity. *American Journal of Psychology*, **69**, 82–6.

Rueckl, J. G. (1990). Similarity effects in word and pseudoword repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 374–91.

Rueckl, J. G. (1993). Jumpnet: A multiple-memory systems connectionist architecture. *Proceedings of the 15th Annual Meeting of the Cognitive Science Society*, pp. 866–71. Hillsdale, NJ: Erlbaum Publishers.

Rueckl, J. G. (submitted). Interference effects in implicit memory. Manuscript submitted for publication.

Rueckl, J. G. and **Dror, I.** (1994). The effect of orthographic-semantic systematicity on the acquisition of new words. In *Attention and performance XV* (ed. C. Umilta and M. Moscovitch), pp. 571–88. Hillsdale, NJ: Erlbaum.

Rueckl, J. G. and **Mathew, S.** (1999). Implicit memory for phonological processes in visual stem completion. *Memory and Cognition*, **27**, 1–11.

Rueckl, J. G. and **Olds, E. M.** (1993). When pseudowords acquire meaning: The effect of semantic associations on pseudoword repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **3**, 515–27.

Rueckl, J. . and **Raveh, M.** (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain and Language*, **68**, 110–7.

Rueckl, J. ., Cave, K. R., and **Kosslyn, S. M.** (1989). Why are 'what' and 'where' processed by separate cortical visual systems? A computational investigation. *Journal of Cognitive Neuroscience*, **1**, 171–86.

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

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

Rumelhart, D. E. and **McClelland, J. L.** (1986b). On learning the past tense of English verbs. In *Parallel distributed processing* (ed. D. E. Rumelhart and J. L. McClelland), Vol. 1, Foundations. Cambridge, MA: MIT Press.

Rumelhart, D., Hinton, G., and **Williams, R.** (1986). Learning internal representations by error propagation. In *Parallel distributed processing: Explorations in the microstructure of cognition* (ed. D. Rumelhart and J. McClelland), Vol. 1, Foundations. Cambridge, MA: MIT Press.

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

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 501–18.

Schacter, D. L. (1992). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. *Journal of Cognitive Neuroscience*, **4**, 244–56.

Schreuder, R. and **Baayen, R. H.** (1997). How complex simplex words can be? *Journal of Memory and Language*, **37**, 118–39.

Seidenberg, M. S. and **McClelland, J. L.** (1989). A distributed, developmental model of visual word recognition. *Psychological Review*, **96**, 523–68.

Seidenberg, M. S., Plaut, D. C., Petersen, A. S., McClelland, J. L., and **McRae, K.** (1994). Nonword pronunciation and models of word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 1177–96.

Sherry, D. and **Schacter, D.** (1987). The evolution of multiple memory systems. *Psychological Review*, **94**, 439–54.

Shevtsova, N. and **Reggia, J. A.** (1999). A neural network model of lateralization during letter identification. *Journal of Cognitive Neuroscience*, **11**, 167–81.

Sloman, S., Hayman, C., Ohta, N., Law, J., and **Tulving, E.** (1988). Forgetting in primed fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 223–39.

Smith, S. M. and **Tindell, D. R.** (1997). Memory blocks in word fragment completion caused by involuntary retrieval of orthographically related primes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23**, 355–70.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, **99**, 195–231.

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

Stark, C. E. L. and **McClelland, J. L.** (2000). Repetition priming of words, pseudowords, and nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **26**, 945–72.

Stolz, J. A. and **Feldman, L. B.** (1995). The role of orthographic and semantic transparency of the base morpheme in morphological processing. In *Morphological aspects of language processing* (ed. L. B. Feldman), pp. 109–30. Hillsdale, NJ: Erlbaum.

Stone, G. O. and **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**, 1248–68.

Tabor, W. and **Tanenhaus, M. K.** (1999). Dynamical models of sentence processing. *Cognitive Science*, **23**, 491–515.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition*, **7**, 263–72.

Taft, M. and **Forster, K. I.** (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, **15**, 607–20.

Tenpenny, P. L. (1995). Abstractionist vs. episodic theories of repetition priming and word identification. *Psychonomic Bulletin and Review*, **2**, 339–63.

Whittlesea, B. W. A. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 3–17.

Zorzi, M., Houghton, G., and **Butterworth, B.** (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 1131–61.

CHAPTER 5

REMI AND ROUSE: QUANTITATIVE MODELS FOR LONG-TERM AND SHORT-TERM PRIMING IN PERCEPTUAL IDENTIFICATION

ERIC-JAN M. WAGENMAKERS, RENÉ ZEELENBERG, DAVE HUBER
JEROEN G. W. RAAIJMAKERS, RICHARD M. SHIFFRIN, AN
LAEL J. SCHOOLER

Abstract

The REM model originally developed for recognition memory (Shiffrin and Steyver 1997) has recently been extended to implicit memory phenomena observed during threshold identification of words. We discuss two REM models based on Bayesian principles: model for long-term priming (REMI; Schooler *et al.*, 2001), and a model for short-term priming (ROUSE; Huber *et al.*, 2001). Although the identification tasks are the same, th basis for priming differs in the two models. In both paradigms we ask whether prior stud merely reflects a *bias* to interpret ambiguous information in a certain manner, or instea leads to *more efficient encoding*. The observation of a 'both-primed benefit' in two-alternative forced-choice paradigms appears to show that *both* processes are present. However, th REMI model illustrates that the both-primed benefit is not necessarily indicative of a increase in perceptual sensitivity but might be generated by a criterion bias. The ROUS model demonstrates how the amount of attention paid to the prime, and the consequen effect upon decision making, may lead to the *reversal* of the normal short-term primin effect that is observed in certain conditions.

Introduction

A stimulus in a current task is said to be *primed* when it has been encountered previously but the memory for that prior occurrence is not required for performance of the curren task. Primed stimuli are generally responded to faster and more accurately than unprimed

RETHINKING IMPLICIT MEMORY

Edited by

JEFFREY S. BOWERS

Department of Experimental psychology
University of Bristol
UK

and

CHAD J. MARSOLEK

Department of Psychology
University of Minnesota
USA

OXFORD
UNIVERSITY PRESS