

# Connectionist Modeling of Relearning and Generalization in Acquired Dyslexic Patients

David C. Plaut

Departments of Psychology and Computer Science  
Center for the Neural Basis of Cognition  
Carnegie Mellon University

October 1997

To appear in J. Grafman and Y. Christen (Eds.), *Neuroplasticity: Building a bridge from the laboratory to the clinic*, New York: Springer-Verlag.

## Abstract

Connectionist models implement cognitive processes in terms of cooperative and competitive interactions among large numbers of simple, neuron-like processing units. Such models provide a useful computational framework in which to explore the nature of normal and impaired cognitive processes. The current work extends the relevance of connectionist modeling in neuropsychology to address issues in cognitive rehabilitation: the degree and speed of recovery through retraining, the extent to which improvement on treated items generalizes to untreated items, and how treated items are selected to maximize this generalization. A network previously shown to model impairments in mapping orthography to semantics was retrained after damage. The degree of relearning and generalization depended on the location of the lesion, and had interesting implications for understanding the nature and variability of recovery in patients. In a second simulation, retraining on words whose semantics are atypical of their category yielded more generalization than retraining on more typical words, suggesting a counterintuitive strategy for selecting items in patient therapy to maximize recovery. Taken together, the findings demonstrate that the nature of relearning in damaged connectionist networks can make important contributions to a theory of rehabilitation in patients.

It was once thought that the brain lost much of its plasticity beyond an early, *critical period* in development. There is now, however, considerable evidence that the response properties of cortical neurons in adult animals can be remapped extensively in response to intensive training regimes (see Kaas, 1994; Merzenich & Jenkins, 1995, for reviews). Although most of these demonstrations have been in the peripheral domains of sensory and motor processing, more recent attempts to apply analogous intervention strategies to language-learning impaired children (Tallal, Miller, & Fitch, 1993) have met with consider-

able success (Merzenich et al., 1996; Tallal et al., 1996). Among other things, these findings raise considerable hope for the development of more effective strategies for remediating the cognitive impairments of individuals with brain damage.

There remains, however, something of a puzzle. If the brain remains so plastic well into adulthood, why is the prognosis for recovery of cognitive functions following brain damage often so poor? While patients with certain types of brain damage may show nearly complete post-morbid recovery (e.g., patients with hemispatial neglect following right parietal damage; see Robertson & Marshall, 1993), others, particularly those with language impairments, do not (see Kertesz, 1985). Moreover, even in the circumstances in which there is substantial recovery of function, little is known about its physiological and cognitive bases, nor what factors might influence its effectiveness (Hillis, 1993).

The current work attempts to provide a theoretical framework, supported by explicit computational simulations, for understanding how and when experience-driven therapy is most effective at remediating cognitive impairments. The focus is not only on the degree and speed with which behavior can be reestablished as a result of therapy, but also on the extent that recovery due to treatment of particular items generalizes to other materials, and the possible bases on which to select items for treatment so as to maximize this generalization. The work is cast in terms of connectionist or parallel distributed processing models, in which information is represented as patterns of activity over large groups of simple, neuron-like units. Processing takes the form of cooperative and competitive interactions among the units on the basis of weighted connections between them. These weights encode the long-term knowledge of the system and are learned gradually through experience in the domain. Models of this form to be devel-

oped within a wide range of cognitive domains, including high-level vision and attention, learning and memory, speech and language processing, and the coordination and control of action (see McClelland, Rumelhart, & the PDP Research Group, 1986; Quinlan, 1991; Rumelhart, McClelland, & the PDP Research Group, 1986b).

The effects of damage in connectionist models have been used to account for a number of specific neuropsychological disorders. Brain damage can be approximated within such models by the removal of some proportion of the units and/or connections in certain regions of the model. Perhaps the most widely investigated class of disorders concern selective impairments in reading—the acquired dyslexias (Hinton & Shallice, 1991; Mozer & Behrmann, 1990; Plaut & Shallice, 1993; Plaut, McClelland, Seidenberg, & Patterson, 1996). The current work extends the relevance of connectionist modeling in cognitive neuropsychology by demonstrating that the same computational principles which are effective for understanding normal cognitive processing, and the effects of brain damage, can also provide insight into the nature of recovery from brain damage.

The next section provides a brief overview of findings from empirical studies attempting to remediate the reading impairments of acquired dyslexic patients. Following this, two computational simulations are presented, both involving networks that are trained to derive the meanings of written words (see Plaut, 1996, for more details). The first demonstrated that, in retraining a network after damage, the degree of relearning and generalization depended on the location of the lesion. The results have interesting implications for understanding the nature and variability of recovery in patients. In the second simulation, retraining on words whose semantics are atypical of their category yielded more generalization than retraining on more typical words, suggesting a counterintuitive strategy for selecting items in patient therapy to maximize recovery. Taken together, the findings demonstrate that the nature of relearning in damaged connectionist networks can make important contributions to a theory of rehabilitation in patients.

## Remediation of Acquired Dyslexia

Coltheart and Byng (1989) undertook a series of remediation studies with a surface dyslexic patient, EE, with left temporal-parietal damage due to a fall. On the basis of a number of preliminary tests, Coltheart and Byng determined that EE had a specific deficit in deriving semantics from orthography. In one study, they gave EE 485 high-frequency words for oral reading. The 54 words he misread were divided in half randomly into treated and untreated sets. For words in the treated set, EE studied cards of the written words augmented with mnemonics for their

meanings. As a result, his reading performance on the treated words improved from 44% to 100% correct. Surprisingly, the untreated words also improved, from 44% to 85% correct. This improvement was not due to “spontaneous recovery” nor to other non-specific effects because performance on the words was stable both before and after therapy. Two other studies with EE produced broadly similar results. Overall, Coltheart and Byng found excellent recovery of treated items and substantial generalization to untreated items (also see Weekes & Coltheart, 1996).

A useful measure of generalization is the amount that untreated items improve relative to the amount that they would have improved if they had been treated directly. This measure can be approximated by the ratio of the improvement on untreated items to the improvement on the treated items. Thus, Coltheart and Byng’s (1989) therapy with EE produced  $41/56 = 73\%$  generalization.

Unfortunately, such promising results are not always found in rehabilitation studies, even those with very similar types of patients. Scott and Byng (1989) treated a surface dyslexic patient for homophone confusions in reading (e.g., TAIL/TALE) and produced improvement on treated items and, to a lesser extent, untreated items, but found no generalization to his writing of the same items (also see Behrmann, 1987). Behrmann and Lieberthal (1989) trained a globally aphasic patient with semantic impairments on a semantic category sorting task. They found improvement on untreated items only within some categories and minimal generalization to items in untreated categories. Finally, Hillis (1993) carried out an extensive rehabilitation program with a patient who had both orthographic and semantic impairments. The patient was able to learn trained tasks (e.g., lexical decision, naming) but showed virtually no generalization to untrained tasks.

Why some patients improve while others do not is not entirely clear. Furthermore, even in those patients who do improve and show generalization, the cause of this generalization—in terms of changes to the underlying cognitive mechanism induced by treatment—is unknown. An explanation of these findings should account not only for the occurrence of generalization in some patients and conditions, but also for its absence in others. As Hillis (1993) points out, what is needed is a theory of rehabilitation that provides a detailed specification of the impaired cognitive system, how it changes in response to treatment, and what factors are relevant to the efficacy of the treatment.

## A Connectionist Approach to Remediation

Early connectionist research (Hinton & Plaut, 1987; Hinton & Sejnowski, 1986) demonstrated that simple net-

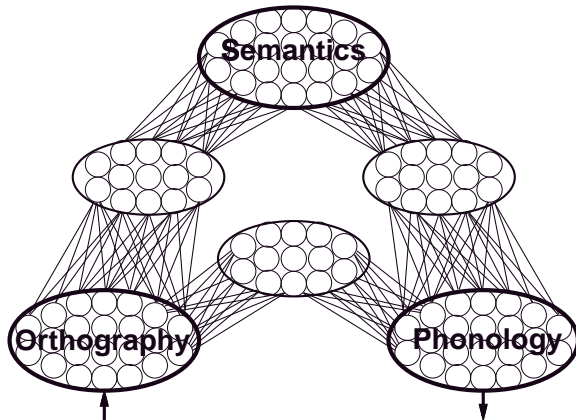


Figure 1. A connectionist framework for lexical processing. Adapted from Plaut (1997).

works trained on unstructured tasks can, when retrained after damage, exhibit rapid recovery on treated items and generalization to untreated items. Plaut (1996) extended these findings to apply directly to understanding the basis and variability of recovery in patients, and to provide a platform for testing hypotheses on how to select items for treatment to maximize generalized recovery. The modeling work was cast within a more general framework of lexical processing (see Seidenberg & McClelland, 1989; Plaut et al., 1996) in which distributed representations of written words (orthography), spoken words (phonology) and their meanings (semantics) interact to simultaneously settle on the best interpretation of an input (see Figure 1).

### Simulation 1: Recovery and Generalization

The network used in the first simulation, depicted in Figure 2, was based closely on the one used by Hinton and Shallice (1991) and constitutes an implementation of the orthography-to-semantics portion of the general framework in Figure 1. Written input, in the form of 40 four-letter words, was presented to the network by clamping particular patterns of activity (1s and 0s) over 32 *orthographic* units (8 features per letter). The meaning of each of the words, falling into 5 categories of 8 items each, was represented by a particular pattern of activity over 68 *semantic* units. The assignment of semantic features to words had the property that words within the same category tended to have more similar semantic representations than those in different categories, and there was no systematic relationship between orthographic similarity and semantic similarity (see Plaut & Shallice, 1993, for details). During processing, the activation levels of units were a smooth, nonlinear (sigmoidal) function of their summed weighted inputs from other units, ranging between 0 and 1. The network was trained with a version of back-propagation appropriate for recurrent net-

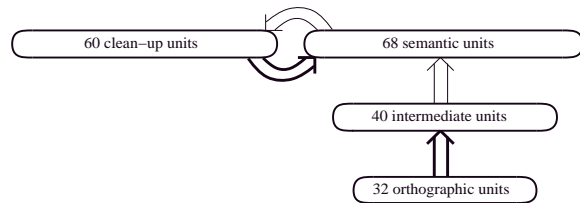


Figure 2. The connectionist network used in the Simulation 1. Arrows represent 25% connectivity between groups of units. The simulation contrasts the effects of lesions to the two sets of connections shown in bold. Adapted from Plaut (1996).

works, known as *back-propagation through time* (Rumelhart, Hinton, & Williams, 1986a), to activate the appropriate semantic features of a word when presented with its orthographic representation. In doing so, the network learned to use bidirectional interactions between the semantic units and so-called *clean-up* units to make the meanings of words into stable, *attractor* patterns.

Once the network had learned to derive the meanings of the 40 words from their written form accurately, it was damaged in one of two places: near orthography (the connections from the orthographic units to the intermediate units) or near semantics (the connections from the clean-up units to the semantic units). These two locations are indicated in bold in Figure 2. Lesions involved removing a random proportion of the indicated set of connections—the severity of the lesion was controlled by changing the proportion of connections that were removed.

After a given instance of lesion, the network was presented with each of the 40 words for processing. As a result of the damage, the semantic activity produced by the network would often differ significantly from the correct semantics of the presented word. The network was considered to have responded correctly if the *proximity* (i.e., normalized dot-product) of the semantics generated by the network was within 0.8 of the correct semantics of the presented word, and the proximity of the next best word was at least 0.05 further (see Hinton & Shallice, 1991, for details). If the generated semantics satisfied these criteria when compared with the semantics of some word other than the one presented, that word was considered to be the network's response (an error). Otherwise, the network was considered to have failed to respond (an omission). The response criteria can be thought of as substituting for the semantics-to-phonology portion of the general framework (see Plaut & Shallice, 1993, for implementations).

For each of the two locations of damage, a severities of lesion was chosen that reduced correct performance on the 40 trained words to 20%. This turned out to involve removing 30% of connections for the orthographic lesions and 50% of connections for the semantic lesions.

Once the performance of the lesioned network on all 40 words was determined, half of the correct words and half of the incorrect words were randomly selected and placed in the *treated* set; the remaining words were placed in the *untreated* set. Thus, both the treated and untreated sets contained 20 words and were balanced for correct performance. For the purpose of setting up the treated and untreated sets, explicit errors and omissions were both considered incorrect and were not distinguished. The lesioned network was then retrained for 50 epochs (training presentations) on the treated words only. Performance was measured at each epoch during relearning separately for the treated and untreated word sets, in terms of the number of words read correctly, and the average proximity of the generated and correct semantics. To ensure that any relearning effects were not simply due to an imbalance in initial performance between the treated and untreated sets, the two sets were exchanged and the retraining was repeated, starting from the same initial set of weights. Thus, each group of words served both as the treated set and the untreated set. Finally, for purposes of comparison, the weights were again reinitialized and the lesioned network was retrained on all 40 words. Results were averaged over twenty instances of each location and severity of lesion, in which a different random subset of connections were removed.

Figure 3 shows the improvement in performance during retraining following lesions near semantics versus lesions near orthography. Considering the former first (see Figure 3a), the network shows rapid relearning of the 20 treated words after a semantic lesion, reaching near perfect performance (98.4% correct) after 50 training presentations. Moreover, performance on the 20 untreated words in this condition also improve considerably, reaching 67.6% correct at this point. The improvement on untreated words was 61% as large as the improvement on the treated words themselves, which is comparable to the 73% generalization found by Coltheart and Byng (1989) for patient EE.

By contrast, retraining after a lesion near orthography produces quite different results (see Figure 3a). In particular, recovery of the 20 treated items is far less effective, although it does reach 93.3% correct after 50 epochs. More critically, performance on the untreated words fails to improve at all, remaining at near 20% correct. Thus, the network exhibited poorer recovery and no generalization following retraining after orthographic lesions.

A clue to the basis for this difference can be found from examining the lesioned network's performance when retrained on all 40 words. Following semantic lesions, recovery is faster when retraining on all 40 words than when retraining on only 20 words. By contrast, following orthographic lesions, retraining on all 40 words produces much poorer recovery than when retraining on only 20 words.

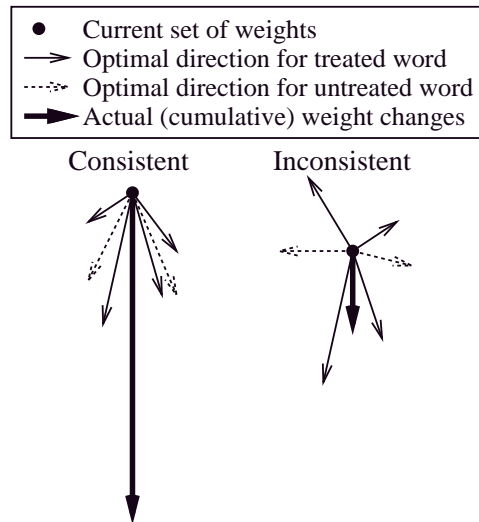


Figure 4. Depiction of the effect of consistent vs. inconsistent weight changes on the extent of recovery and generalization in relearning. In each condition, the small solid arrows represent directions of weight change induced by treated words; the large solid arrow is the (vector) sum of these smaller arrows, representing the actual weight changes administered to the network. The length of this vector reflects the speed of relearning the treated words. The dotted arrows represent directions of weight change that would be optimal for untreated words if they were trained—to the extent that these point in the same direction as the actual weight change vector, retraining on the treated words will also improve performance on the untreated words. Adapted from Plaut (1996).

These findings make sense if the weight changes induced by retrained words after semantic lesions are more consistent across words than after orthographic lesions. The actual weight changes administered to the network after a retraining epoch are the sum of the weight changes induced by each individual word (scaled by the learning rate). Weight changes that are consistent across retrained words accumulate, resulting in fast learning; weight changes that are inconsistent cancel each other out, resulting in much slower learning. Figure 4 presents a graphical depiction of this effect using vectors (arrows) to represent weight changes. Within semantics, similar words require similar interactions, so that the weight changes caused by retraining on some words will tend also to improve performance on other, related words (i.e., the optimal weight changes for words are mutually consistent). By contrast, similar orthographic patterns typically must generate very different semantic patterns. As a result, when retraining after lesions near orthography, the weight changes for treated items are unrelated to those that would improve the untreated items, and there is no generalization.

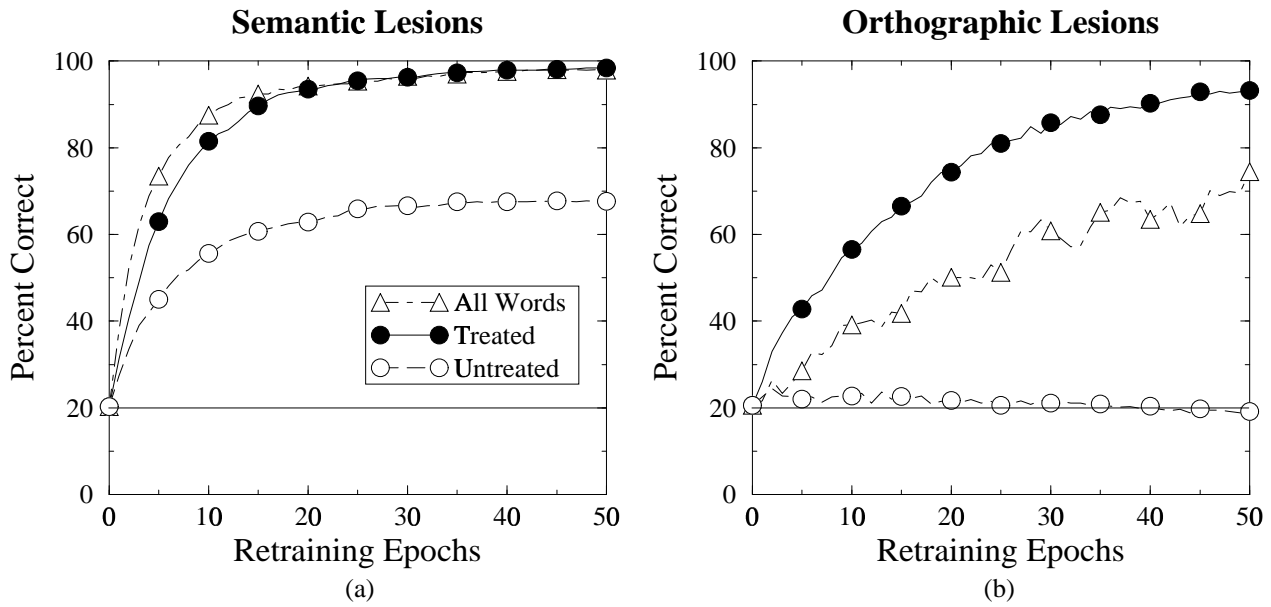


Figure 3. Improvement on treated and untreated items when retraining a network that maps orthography to semantics after (a) lesions near semantics, or (b) lesions near orthography. Adapted from Plaut (1996).

As reviewed earlier, studies of cognitive rehabilitation of acquired dyslexics have demonstrated considerable relearning of treated items and (often) improvement on untreated but related items. At a general level, the cause of rapid relearning and generalization in the network may provide an explanation for the nature of recovery in these patients. At a more specific level, the finding that the extent of relearning depends on the location of damage may provide an explanation for why only some patients show substantial recovery and generalization. The simulation results suggest that a patient with a functional impairment close to or within semantics should show considerable generalization, while one with an impairment close to orthography should show little or none. Conversely, the degree of generalization observed in a patient can be used to predict the fine-grained location of their functional impairment *within* the semantic route.

### Simulation 2: Designing Retraining to Maximize Generalization

Ideally, an understanding of the impairment in a particular patient should lead to the design of a rehabilitation strategy that maximizes recovery. A potential benefit of connectionist modeling in neuropsychological rehabilitation is that it can provide a framework for investigating the relative effectiveness of alternative rehabilitation strategies. One aspect of a retraining simulation that is under experimental control, and that might influence the nature of recovery, is the selection of items for treatment.

An important aspect of the structure of semantic repre-

sentations, at least of nouns, is that they are organized into categories. Furthermore, relative to this category structure, a critical semantic variable is typicality—how close a concept is to the central tendency of its category (Rosch, 1975). For instance, a robin is highly typical among birds, an eagle is less typical, and a penguin is highly atypical. The question is, is it better to retrain on typical or atypical words? A natural intuition is that relearning the central tendency of a category—that is, retraining on typical words—should lead to the greatest generalization to other words in the category. The results of the current simulation, however, show the opposite: retraining on words that are somewhat atypical of their semantic category leads to greater generalization than retraining on more typical words. The reason, put briefly, is that atypical words collectively convey more information on the overall structure of the category—specifically, on how semantic properties can vary across category members—while still providing a good approximation of the central tendency of the category.

The simulation used an artificial version of the task of mapping orthography to semantics in order to more carefully control the nature of the semantic categories. The training set consists of 100 artificial “words.” The orthographic representation of each word was created by randomly assigning it an average of 4 out of 20 possible orthographic features. The semantic representations of words were generated by distorting a “prototype” pattern which was generated randomly to have 10 of 50 possible semantic features. The degree of typicality of words is reflected in the number of features that its representation dif-

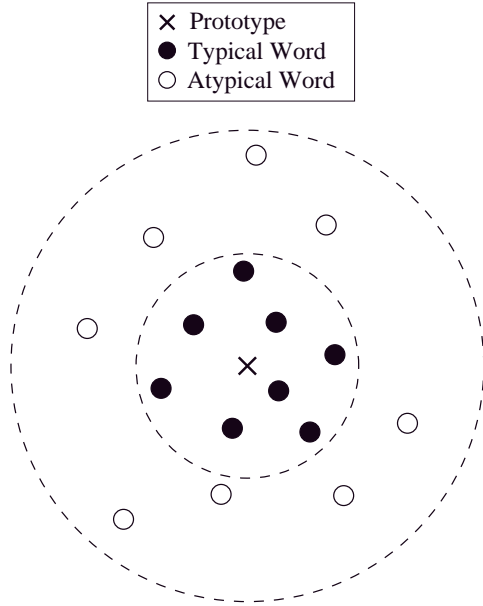


Figure 5. A depiction of the relationship in semantic space between the prototype of a category and typical versus atypical exemplars in that category. Adapted from Plaut (1996).

fers from the prototype—typical words share most of the features of the prototype, while atypical words share far fewer. To implement this, two sets of 50 word meanings were generated from the prototype using different levels of random distortion. The *typical* set consisted of instances produced by a small distortion of the prototype—each semantic feature had a probability  $d = 0.1$  of being randomly regenerated (with probability  $p = 0.2$ ). The *atypical* set consisted of instances generated using a large distortion ( $d = 0.5$ ). Geometrically, if the prototype corresponds to a particular point in the space of semantic representations, the typical words are points that are near the prototype, while the atypical words are farther away (see Figure 5). Orthographic patterns were assigned to semantic patterns randomly to ensure that, as in English, there was no systematic relationship between orthography and semantics.

A network was trained with back-propagation through time to activate the correct semantic pattern for each of the 100 artificial words when presented with its orthographic representation. The architecture of the network was broadly the same as the network from the first simulation (see Figure 2). The differences were that, in the current network, there were only 20 orthographic units (compared with 32), 50 semantic units (compared with 68), and 50% of the possible connections between groups of units are included (compared with 25%).

After training, the network was lesioned by removing

a randomly-selected 25% of the connections between the intermediate units and the semantic units. This lesion location and severity was selected because it produces an intermediate amount generalization (27%; see Plaut, 1996), providing a clear opportunity for the composition of the treated set to have either a positive or negative impact on generalization.

After each lesion, performance on all 100 words was measured. A presented word was considered correct if the semantics generated by the network had a higher proximity (normalized dot product) to the correct semantics for the word than to the semantics for any other word. Based on this initial performance, the typical and atypical word sets each were divided in half, balancing for correct performance. The lesioned network was then retrained for 50 epochs, either on half of the typical words or on half of the atypical words (25 words). During retraining, improvement in correct performance was measured on this treated set as well as on two untreated sets: the remaining words of the same type (typical or atypical), and all of the words of the other type. Each half of each group in turn served as the treated set for retraining (reinitializing the weight each time). In this way, the retraining procedure was able to measure the generalization to typical and atypical words when retraining on typical or atypical words.

Somewhat surprisingly, although retraining on typical exemplars produced greater recovery on treated items, retraining on atypical exemplars produced greater generalization to untreated items (see Figure 6). These findings make sense given the adequacy with which sets of typical versus atypical exemplars approximate the range of semantic similarity among all of the words. Semantically typical words accurately estimate the central tendency of a category, but provide little information about the ways in which category members can vary. By contrast, each atypical word indicates many more ways in which members can differ from the prototype and yet still belong to the category. Thus, collectively, the semantic representations of atypical words cover more of the features needed by the entire set of words than do the representations of more typical words. At the same time, the average effects of retraining on atypical words provides a reasonable estimate of the central tendency of the category, yielding generalization to typical words (as found in human category learning by, e.g., Posner & Keele, 1968). In this way, the simulation generated a novel prediction about how to select items for treatment so as to maximize generalized recovery.

## Conclusion

Attempts at cognitive rehabilitation of acquired dyslexic patients have resulted in considerable improvement in per-

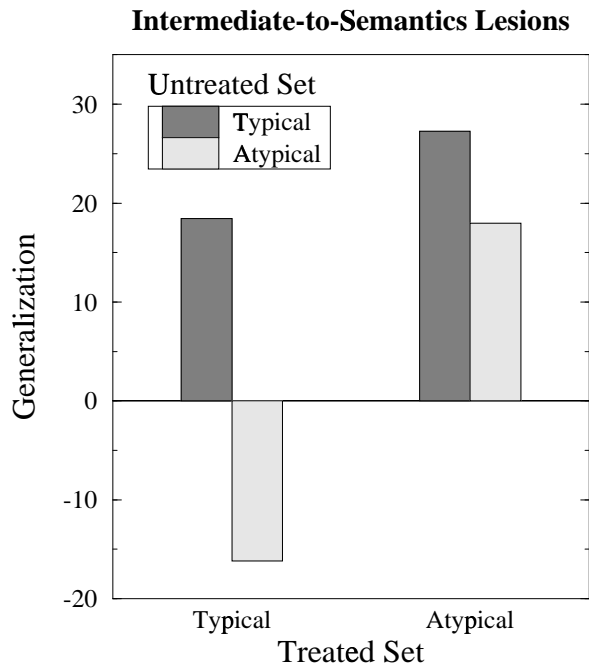


Figure 6. Generalization from retraining after lesions of 25% of the intermediate-to-semantic connections, as a function of the semantic typicality of the treated and untreated sets. Adapted from Plaut (1996).

formance on treated words, as well as significant generalization to untreated but related words, although the degree of recovery across patients can vary considerably. There is, however, little understanding of the underlying mechanisms by which cognitive functions recover, either spontaneously or as a direct result of therapeutic intervention. Generalization in the domain of reading via meaning is particularly puzzling as there is no systematic relationship between the written or spoken forms of words and their meanings.

Connectionist modeling offers specific hypotheses about the nature of the representations and computations that underlie cognitive processes, as well as how these processes are learned through experience and how they are affected by brain damage. The current work attempts to extend the relevance of connectionist modeling in neuropsychology one step further, to contribute to a theory of rehabilitation (see Hillis, 1993) based on analyses of relearning in damaged networks. To this end, two simulations were carried out in the domain of reading via meaning, to address a central issue in rehabilitation studies: What factors influence the degree of recovery and generalization observed in patients?

A general finding that emerged from the simulations is that the degree of recovery and generalization produced by retraining after damage depends strongly on the relative structure of the tasks performed by different parts

of the system, and the extent to which the items selected for treatment approximate this structure. Specifically, the first simulation found robust recovery and generalization from retraining following a lesion near semantic, but not following a lesion near orthography. In this way, the network results may help explain the variability in recovery observed in patients. The second simulation found, somewhat surprisingly, that retraining on less typical exemplars within a category produced greater generalization than did retraining on more typical exemplars. Overall, the results demonstrate that investigations of relearning after damage in connectionist networks can provide an account of the general nature of relearning and generalization in patients and can generate interesting hypotheses about the design of effective patient therapy.

## References

Behrmann, M. (1987). The rites of righting writing: Homophone remediation in acquired dysgraphia. *Cognitive Neuropsychology*, 4, 365–384.

Behrmann, M., & Lieberthal, T. (1989). Category-specific treatment of a lexical semantic deficit: A single case study of global aphasia. *British Journal of Communication Disorders*, 24, 281–299.

Coltheart, M., & Byng, S. (1989). A treatment for surface dyslexia. In X. Seron, & G. Deloche (Eds.), *Cognitive approaches in neuropsychological rehabilitation* (pp. 159–174). Hillsdale, NJ: Erlbaum.

Hillis, A. E. (1993). The role of models of language processing in rehabilitation of language impairments. *Aphasiology*, 7, 5–26.

Hinton, G. E., & Plaut, D. C. (1987). Using fast weights to deblur old memories. In *Proceedings of the 9th Annual Conference of the Cognitive Science Society* (pp. 177–186). Hillsdale, NJ: Erlbaum.

Hinton, G. E., & Sejnowski, T. J. (1986). Learning and relearning in Boltzmann Machines. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (pp. 282–317). Cambridge, MA: MIT Press.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, 98, 74–95.

Kaas, J. H. (1994). The reorganization of sensory and motor maps in adult mammals. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 51–71). Cambridge, MA: MIT Press.

Kertesz, A. (1985). Recovery and treatment. In K. M. Heilman, & E. Valenstein (Eds.), *Clinical neuropsychology* (pp. 481–505). New York: Oxford University Press.

- McClelland, J. L., Rumelhart, D. E., & the PDP Research Group (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models*. Cambridge, MA: MIT Press.
- Merzenich, M. M., & Jenkins, W. M. (1995). Cortical plasticity, learning and learning dysfunction. In B. Julesz, & I. Kovacs (Eds.), *Maturational windows and adult cortical plasticity* (pp. 247–272). Reading, MA: Addison-Wesley.
- Merzenich, M. M., Jenkins, W. M., Johnson, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science*, *271*, 77–81.
- Mozer, M. C., & Behrmann, M. (1990). On the interaction of selective attention and lexical knowledge: A connectionist account of neglect dyslexia. *Journal of Cognitive Neuroscience*, *2*, 96–123.
- Plaut, D. C. (1996). Relearning after damage in connectionist networks: Toward a theory of rehabilitation. *Brain and Language*, *52*, 25–82.
- Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of naming and lexical decision. *Language and Cognitive Processes*, *12*.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.
- Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*, 377–500.
- Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.
- Quinlan, P. (1991). *Connectionism and psychology: A psychological perspective on new connectionist research*. Chicago: University of Chicago Press.
- Robertson, I. E., & Marshall, J. C. (Eds.). (1993). *Unilateral neglect: Clinical and experimental studies*. Hillsdale, NJ: Erlbaum.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*, 192–233.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986a). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (pp. 318–362). Cambridge, MA: MIT Press.
- Rumelhart, D. E., McClelland, J. L., & the PDP Research Group (Eds.). (1986b). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations*. Cambridge, MA: MIT Press.
- Scott, C., & Byng, S. (1989). Computer assisted remediation of a homophone comprehension disorder in surface dyslexia. *Aphasiology*, *3*, 301–320.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568.
- Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: A case for the preeminence of temporal processing. In P. Tallal, A. M. Galaburda, R. R. Llinas, & E. von Euler (Eds.), *Temporal information processing in the nervous system: Special reference to dyslexia and dysphasia*. (pp. 27–47). New York: New York Academy of Sciences.
- Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagaraja, S. S., Schreiner, C., Jenkins, W. M., & Merzenich, M. M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, *271*, 81–84.
- Weekes, B., & Coltheart, M. (1996). Surface dyslexia and surface dysgraphia: Treatment studies and their theoretical implications. *Cognitive Neuropsychology*, *13*, 277–315.