Perseverative and Semantic Influences on Visual Object Naming Errors in Optic Aphasia: A Connectionist Account

David C. Plaut
Carnegie Mellon University

Tim Shallice
University College London

Abstract

■ Although perseveration—the inappropriate repetition of previous responses—is quite common among patients with neurological damage, relatively few detailed computational accounts of its various forms have been put forth A particularly well-documented variety involves the pattern of errors made by "optic aphasic" patients, who have a selective deficit in naming visually presented objects Based on our previous work in modeling impaired reading via meaning in deep dyslexia, we develop a connectionist simulation of visual object naming The major extension in the present work is the incorporation of short-term correlational weights that bias the network towards reproducing patterns of activity that have occurred on recently preceding trials Under damage, the network replicates the complex semantic and perseverative effects found in the optic aphasic error pattern Further analysis reveals that the perseverative effects are strongest when the lesions are near or within semantics, and are relatively mild when the preceding object evokes no response Like optic aphasics, the network produces predominantly semantic rather than visual errors because, in contrast to reading, there is some structure in the mapping from visual to semantic representations for objects Viewed together with the dyslexia simulations, the replication of complex empirical phenomena concerning impaired visual comprehension based on a small set of general connectionist principles strongly suggests that these principles provide important insights into the nature of semantic processing of visual information and its breakdown following brain damage ■

INTRODUCTION

In neuropsychology, "perseveration" refers to the continuation or repetition of an activity or percept when the eliciting stimulus is no longer present and typically when it has been replaced by a different one It is very commonly observed For example, Helmick and Berg (1976), who used a variety of tasks involving naming, drawing, defining, and the like, found that responses on 10% of trials for a group of 30 brain-damaged patients were some form of perseveration

Virtually the entire gamut of behaviors tested in neurological patients can give rise to perseverations At the more peripheral end of the range are visual perseverations (palinopsias), most frequently found after right occipital or parieto-occipital lesions (Hécaen & Angelergues, 1963) They involve the reappearance of percepts when the object that gave rise to them is no longer present By contrast, at the central end of the range are what Sandson and Albert (1984) call "stuck-in-set" perseverations observed after frontal lesions in tasks such as Wisconsin Card Sorting, where the patient has to switch from sorting cards by one rule (e g , shape) to sorting by another (e g , number) without an instruction as to which rule has become the correct one On this task, frontal patients tend to continue sorting by the previously correct rule (Milner, 1963)

Stuck-in-set perseveration has recently been analyzed from a theoretical perspective (see Dehaene & Changeux, 1991; Levine & Prueitt, 1989; Shallice, 1982) However, as Sandson and Albert (1984) have pointed out, the sets of behaviors termed "perseverative" are functionally heterogeneous No other type of perseveration has received a computational analysis In this paper, we analyze a second major subtype of perseveration, called by Sandson and Albert "recurrent perseveration," which they hold to be related anatomically to posterior left hemisphere damage It is the unintentional repetition of a previous response to a subsequent stimulus Thus, the patient described by Lhermitte and Beauvois (1973) named a comb correctly, but then named a fork that was presented on the subsequent trial as a comb as well

There are relatively few detailed empirical analyses of the performance of patients showing strong persevera-

tive tendencies of this type However, one group of disorders provides a clear exception and is therefore a good candidate for more detailed modeling These are the selective problems in naming stimuli presented in one particular modality, of which an example was given above They can arise after posterior lesions to the left hemisphere

A common way of classifying such disorders where the problem is specific to the visual modality is to differentiate between associative visual agnosia and optic aphasia, with the former being a difficulty in recognizing visually presented objects and the latter a difficulty in naming them (see, e g , Beauvois, 1982; Farah, 1990; McCarthy & Warrington, 1990) There are, however, a variety of views on the relation between these disorders (see, e g , Beauvois, 1982; Humphreys & Riddoch, 1987) and, indeed, on the relation between associative agnosia and the disorder with which it is standardly contrasted, namely apperceptive agnosia (see, e g , Farah, 1990; McCarthy & Warrington, 1990) Quite a number of authors see no substantial difference between associative agnosia and optic aphasia, but they disagree as to which is primary Thus, some view optic aphasia as being merely a subtle form of agnosia (Goodglass, Barton, & Kaplan, 1968; Rubens, 1979) Other authors, however, argue that visual perception is intact in optic aphasics Indeed, such patients can often mime appropriately the use of an object they cannot name (Gil, Pluchon, Toullat, Michenau, Rogez, & Levevre, 1985; Lhermitte & Beauvois, 1973; Riddoch & Humphreys, 1987) but debate continues as to whether satisfactory performance of such tasks requires identification of the object, or whether they can be carried out by some mix of superordinate knowledge and structural descriptions (see Farah, 1990; Riddoch & Humphreys, 1987; Shallice, 1988a) Finally, some authors deny the autonomy of associative agnosia, viewing it as either a subtle apperceptive agnosia (e g , Farah, 1990) or an optic aphasia (e g , Geshwind, 1965)

Whether or not associative agnosia and optic aphasia should be distinguished, a recent review (Iorio, Falanga, Fragassi, & Grossi, 1992) shows that patients with posterior left hemisphere lesions who have been categorized as exhibiting one or the other syndrome produce very similar patterns of errors (see Table 1) Typically, they produce semantic errors (e g , *needle* → "thread"), perseverative errors (*needle* → "cat" after naming a *cat*), and unrelated errors (e g , *needle* → "house") We will call this the "optic aphasic" error pattern Some also produce what Iorio and colleagues call "confabulations," but which Beauvois (1982) describes as "a peculiar mingling of semantic errors, perseverations, and descriptions of morphological features of the object" Visual errors (e g , *needle* → "toothpick") are less frequently recorded By contrast, almost all patients with bilateral lesions who were clinically diagnosed as associative agnosic produced predominantly visual errors

In this paper we will be concerned with modeling perseveration in the visual and semantic processes beyond the level of structural descriptions In general, associative agnosic and optic aphasic patients appear relevant Some authors, however, have used the term "associative agnosia" to encompass patients with subtle problems at the structural description level (see Farah, 1990, for discussion) In addition, agnosic problems arising from right hemisphere or bilateral lesions are typically assumed to be at or before that level (see, e g , Humphreys & Riddoch, 1987; McCarthy & Warrington, 1986) Therefore, for a relevant patient database on which to test a model of the appropriate levels of perceptual and semantic processing, we will restrict consideration to associative agnosics with left hemisphere lesions and optic aphasic patients, who all have left hemisphere lesions [1]

As far as these left posterior patients are concerned, there are only two clear exceptions who do not show the optic aphasic error pattern: associative agnosic FRA (McCarthy & Warrington, 1986), who made only superordinate errors, and an optic aphasic patient described by Coslett and Saffran (1989), who made unrelated errors almost exclusively These patients will be considered in the Discussion

The most detailed account of a patient with a specific visual naming difficulty who exhibited a semantic and perseverative error pattern was provided by Lhermitte and Beauvois (1973) for their patient JF, who had had a left posterior cerebral artery infarct JF was virtually normal at naming from verbal description (4% errors) and from touch (8% errors) but made many more errors at

Table 1. Cases Reviewed by Iorio et al (1992) for Whom the Types of Naming Errors Are Reported, Subdivided According to Diagnostic Category and Location of Lesion

		Error Type			
	Cases	*Visual*	*Semantic*	*Perseverative*	*Confabulations*
Bilateral considered "visual associative agnosic"	10	90	20	10	0
Unilateral (left) considered "visual associative agnosic"	8	42	88	88	50
"Optic aphasic"	10	40	100	80	40

naming pictures (28% errors) and objects (23%) Lhermitte and Beauvois (1973) identified two apparent influences on the error responses—"horizontal" ones directly from the stimulus, and "vertical" ones from an earlier stimulus (i e , errors showing some perseverative effect) Table 2 shows the influences apparent in errors made by JF on naming 100 pictures and 30 objects In a further experiment, where 320 color and black-and-white pictures were presented and 176 errors were made, the rate of errors with a vertical influence remained high (over 50% of all errors) and varied little with picture type These perseverations were mainly mixed with other types of error The ratio of semantic errors to visual errors varied with the type of picture, being greater with more complex colored pictures than with black-and-white sketches

Given the existence of such a well-documented patient who shows the optic aphasic error pattern, together with the variety of descriptions of other patients who have qualitatively similar error patterns, it is appropriate to consider how the errors arise There has, however, been virtually no discussion of this issue Riddoch and Humphreys (1987) discuss the relevance of the high rate of mixed visual-and-semantic errors in their optic aphasic patient, JB, but as he made few perseverative errors they do not discuss that critical aspect of the symptom pattern

By contrast, most theoretical discussion has focused on the overall pattern of performance exhibited by optic aphasic and associative agnosic patients At least two explanations have been offered for associative agnosia One is degradation of the representations in a partially separable "visual" semantic system (e g , McCarthy and Warrington, 1986) The other is the existence of a subtle impairment in visual perceptual processing (Farah, 1990)

Table 2. Errors Made by JF in Two Experiments[a]

	100 Pictures	*30 Objects*
Horizontal errors		
Semantic		
(e g , *shoe* → "hat")	9	3
Visual		
(e g , *coffee beans* → "hazel nuts")	2	1
Mixed visual-and-semantic		
(e g , *orange* → "lemon")	6	1
Vertical errors		
Item and coordinate perseveration		
(e g , T26 → "wristwatch"		
T27 *scissors* → "wristwatch"		
(e g , T44 → "newspaper"		
T45 *case* → "two books")	8	2
Mixed horizontal/vertical errors		
(e g , T43 → "chair"		
T47 *basket* → "cane chair"		
T53 *string* → "strand of weaved cane")	3	0

[a]From Lhermitte and Beauvois (1973)

Five possible accounts have been suggested for optic aphasia Ratcliff and Newcombe (1982) have argued that there is a "direct" route from visual perceptual processing to naming, analogous to the lexical nonsemantic route from orthographic to phonological processing in reading (Morton & Patterson, 1980; Schwartz, Saffran, & Marin, 1980) However, there is no independent evidence for the existence of such a route (see Howard & Franklin, 1988)

A second explanation (e g , Beauvois, 1982) is that semantics is not a unitary entity, but is separated into "visual" and "verbal" components Visual input can directly access only visual semantics, and naming can be based only on verbal semantics Visual object naming requires communication from visual to verbal semantics; optic aphasia arises naturally from a disconnection between them Intact gesturing and categorization can be based on visual semantics, while intact auditory recognition is based on direct access to verbal semantics The main problem with this account concerns the adequacy of the concept "visual semantics" (for discussion, see Caramazza, Hillis, Rapp, & Romani, 1990; Riddoch, Humphreys, Coltheart, & Funnell, 1988; Shallice, 1988a; 1993)

A related proposal (Coslett & Saffran, 1989) is that semantics is divided not by modality but by *hemisphere,* with naming only supported in the left hemisphere On this hypothesis, optic aphasia arises when visual input from both hemispheres is disconnected from left-hemisphere semantics, with residual comprehension subserved by right-hemisphere semantics In essence, this theory parallels the right-hemisphere hypothesis for reading in deep dyslexia (Coltheart, 1980; 1983; Saffran, Boygo, Schwartz, & Marin, 1980) and shares many of its strengths and weaknesses (see Coltheart, Patterson, & Marshall, 1987; Patterson & Besner, 1984; Shallice, 1988a)

Yet another account of optic aphasia (Gordon, 1982; Margolin, Marcel, & Carlson, 1985; Rubens, 1979) locates the impairment to the transmission of information from visual perceptual processing to a unitary semantics, and challenges the claim that recognition is intact in these patients On this approach, optic aphasia would amount to a type of "semantic access agnosia" (Riddoch & Humphreys, 1987) We return to the issue of the intactness of recognition in optic aphasic patients in the Discussion

A final, more recent proposal (Farah, 1990) hypothesizes that optic aphasics have two partial lesions, one between vision and semantics, and the other between semantics and naming Each separate impairment is sufficiently mild to allow reasonable performance on gesturing or tactile naming, but tasks that require both pathways—visual naming—would be much more dras-

tically impaired The explanation for the disproportionate naming impairment hinges on the notion that an interactive connectionist system would be sufficiently robust under each partial lesion alone, but would show *superadditive* impairment under the combined lesions. Unfortunately, preliminary simulations exploring this possibility, carried out by one of us (DP) in collaboration with M Farah, have been unsuccessful to date

None of these proposals for explaining the overall pattern of performance across tasks in optic aphasia is completely satisfactory—each involves either ad hoc assumptions or insufficiently supported claims From the present perspective they have an even more serious disadvantage—they all focus on explaining the dissociation between impaired naming and (relatively) intact recognition, but have little to say about other characteristics of the disorder—in particular, the types of errors the patients make in naming

In the present paper we will therefore reverse the standard explanatory emphasis We will be concerned with providing an account of the error pattern exhibited in optic aphasia Any dissociation between naming and recognition (as demonstrated, for example, by mime) will be treated as a secondary issue The error pattern, as pointed out earlier, is actually found not only in optic aphasic patients, but also in most left hemisphere patients described as associative agnosic We refer to it as the optic aphasic error pattern for simplicity Possible differences between these disorders in the overall pattern of performance will be ignored until the Discussion, and an explanation of the error pattern that might apply to both disorders will be sought

The approach we will take is most closely related to the theoretical position of Riddoch and Humphreys (1987) in that we will assume that the error pattern arises from a lesion that impairs the generation of semantics from visual structural descriptions, and that this process operates on a "cascade" principle (i e , levels continually pass on partial information throughout their computation—see McClelland, 1979)[2] However, to explain the error pattern we need to assume that the visual-to-semantics process can be effectively modeled in connectionist terms Support for this approach comes from our recent work in modeling reading via meaning in patients with "deep dyslexia" (Hinton & Shallice, 1991; Plaut & Shallice, 1991a,b; 1993)

CONNECTIONIST MODELING OF IMPAIRED VISUAL COMPREHENSION

Deep dyslexic patients can pronounce written words only by first accessing their meaning, and typically make errors in this process (e g , misreading the word RIVER as "ocean") In addition to these "semantic" errors, a number of other error types occur: visual (e g , SWORD → "words"), derivational (e g , GROWN → "growing"), visual-and-semantic (e.g., TROUBLE → "terrible"), and visual-then-semantic (e g , SYMPATHY → "orchestra," presumably via *symphony*) Furthermore, these patients exhibit a variety of other symptoms, including an effect of part-of-speech in correct performance (nouns > adjectives > verbs > function words), better performance on concrete, highly imageable words than on abstract, less imageable words, and an inability to read pronounceable nonwords

Through a series of simulation experiments carried out in collaboration with G Hinton (Hinton & Shallice, 1991; Plaut & Shallice, 1991a,b; 1993), we have demonstrated that the characteristics of deep dyslexia arise when systems of a particular type are lesioned in virtually any part The systems are those that have the following characteristics:

1 Orthographic and semantic representations are distributed over separate groups of units, such that similar patterns represent similar words in each domain, but similarity is unrelated between domains

2 Connection weights are learned by a procedure for performing gradient descent in some measure of performance on the task of mapping orthography to semantics

3 Mapping orthography to semantics is accomplished through the operation of "attractors "

4 The semantic representations of concrete words are much "richer" than those of abstract words (i e , contain considerably more consistently accessed features)

Figure 1 shows the architecture of the network used by Hinton and Shallice (1991) Orthography is represented in terms of four groups of "grapheme" units, in which each unit represents a particular letter at a particular position within the word (McClelland & Rumelhart, 1981) The semantics of each word is described in terms of predetermined semantic features designed to capture intuitive semantic distinctions The network has two main pathways: (1) a "direct" pathway, from grapheme units to semantic units via intermediate units, and (2) a "clean-

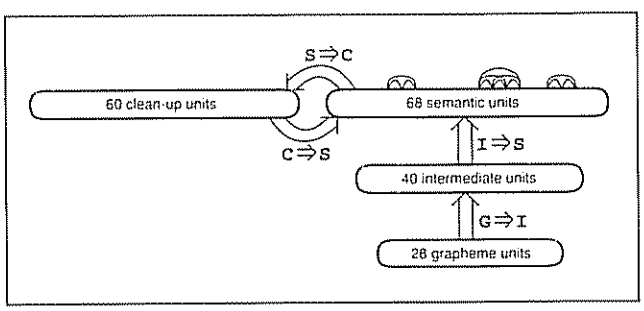

Figure 1. The network used by Hinton and Shallice (1991) Arrows represent sets of connections that are lesioned Notice that sets of connections are labeled with the initials of the names of the source and destination unit groups (e g G→I for *grapheme-to-intermediate* connections) Only 25% of the possible connections in each pathway were included However, additional direct connections were added among semantic units that represent closely related features

up" pathway, from the semantic units to clean-up units and back to the semantic units The direct pathway generates initial semantic activity from visual input The clean-up pathway iteratively refines this initial activity into the exact correct semantics of the word

The network was initialized to have small random weights, so that at the beginning of training the pattern of semantic activity produced by the word was quite different from its correct semantics An iterative version of the backpropagation learning procedure, known as "backpropagation through time" (Rumelhart, Hinton, & Williams, 1986; Williams & Peng, 1990, see Appendix A), was used to train the network to activate each semantic unit to within 0 1 of its correct value for each word

After training, the network was systematically "lesioned" by either removing some units or connections, or by adding random noise to the weights As a result of this damage, the semantics produced by the network typically differs somewhat from the exact semantics of the presented word Hinton and Shallice (1991) defined certain criteria that had to be satisfied by the corrupted semantics in order for the damaged network to be considered to have made a response Figure 2 shows the distribution of error types for *disconnection* lesions of each main set of connections Lesions throughout the network reproduce the cooccurrence of semantic, visual, and mixed visual-and-semantic errors found in deep dyslexia

More recently, we have extended these initial findings in two ways (Plaut & Shallice, 1991a,b; 1993) First, we have established the generality of the deep dyslexic error pattern by showing that it does not depend on peculiar characteristics of the network architecture, the learning procedure, or the way responses are generated from semantic activity Second, we have extended the approach to account for many of the remaining characteristics of deep dyslexia, including the effects of concreteness/imageability and their interaction with visual errors, the occurrence of visual-then-semantic errors, greater confidence in visual as compared with semantic errors, relatively preserved lexical decision with impaired naming, and the existence of different subvarieties of deep dyslexia

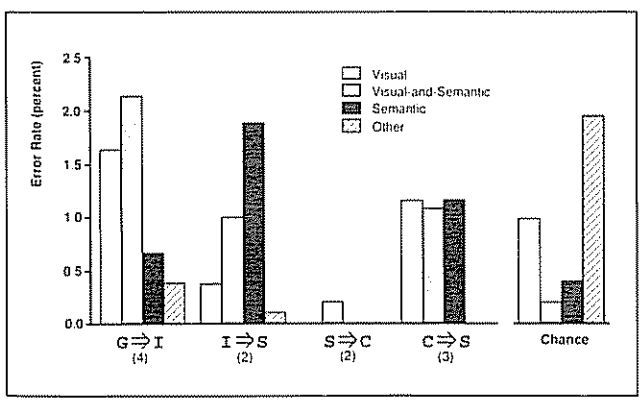

Figure 2. The distribution of error types produced by *disconnection* lesions that resulted in 25–75% correct performance in the model, as well as the distribution of types for error responses chosen randomly from the word set The absolute height of the "Chance" distribution is arbitrary—only the relative rates are informative The number of lesion severities included in the calculation of error rates is indicated in parentheses below the label for each lesion location

A critical concept in understanding these results is that of an "attractor" The semantic units in the network change their states over time in response to a particular orthographic input The initial pattern of semantic activity generated by the direct pathway may be quite different from the exact semantics of the word Interactions among semantic units, either directly via connections among themselves, or indirectly via the clean-up units, serve to gradually modify and "clean-up" the initial pattern into the final, correct pattern This process can be conceptualized in terms of movement in the high-dimensional space of possible semantic representations, in which the state of each semantic unit is represented along a separate dimension At any instant in processing a word, the entire pattern of activity over the semantic units corresponds to a particular point in semantic space The exact meanings of familiar words correspond to other points in the space The states of semantic units change over time in such a way that the point representing the current pattern of semantic activity "moves" to the point representing the nearest familiar meaning In other words, the pattern corresponding to each known word meaning becomes an "attractor" in the space of semantic representations: patterns for nearby but unfamiliar meanings move toward the exact pattern of the nearest known meaning The region in semantic space corresponding to the set of initial patterns that move to a given attractor is called its "basin" of attraction As a result of damage to the network, the initial semantic pattern generated by a word may be "captured" within the basin of a related word (see Fig 3) Since the layout of attractor basins must be sensitive to both visual and semantic similarity, these metrics are reflected in the types of errors that occur as a result of damage

Our ability to account for a number of aspects of deep dyslexia, and in particular the cooccurrence of a number of different types of error, using a connectionist model based on attractors suggests that a related approach may be possible for another neuropsychological error pattern, that of optic aphasia

SHORT-TERM CORRELATION WEIGHTS

The cooccurrence of visual, semantic, and mixed visual-and-semantic errors in deep dyslexic reading would appear to be analogous to the optic aphasic error pattern in object naming, suggesting a natural account of the latter in terms of a network that maps visual representations onto semantic representations using attractors However, the perseverative effects in optic aphasia, and

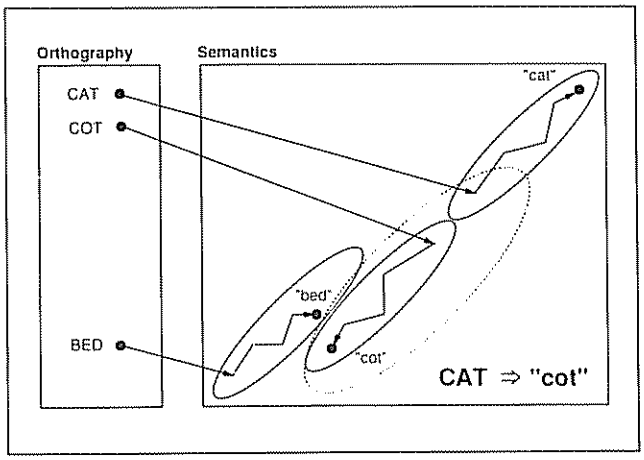

Figure 3. How damage to semantic attractors can cause a visual error Points in each space correspond to particular patterns of activity over the corresponding group of units The solid ovals depict the normal basins of attraction; the dotted one depicts a basin after semantic damage

their interactions with semantic effects, are less straightforward In the dyslexia simulations, the networks were completely reset before the presentation of each word—there was no opportunity for the response to one stimulus to influence responses to subsequent stimuli Accounting for the perseverative effects in optic aphasia requires an elaboration of the computational formalism

There are many possible ways of introducing effects of the temporal order of stimulus presentation into connectionist networks Perhaps the simplest would be to process each object beginning from the set of unit states corresponding to the interpretation of the previous object, rather than resetting the network However, perseverative effects in optic aphasics can span intervening objects (Lhermitte & Beauvois, 1973), which would be difficult to account for solely in terms of sustained activity across object presentations (see also Joordens & Besner, 1992)

The approach we adopt involves introducing short-term weights that depend on the recent correlations between unit states These weights augment the standard long-term weights that are slowly modified over the course of learning In particular, each connection is given a second weight whose value is an exponentially decaying weighted average over stimulus presentations of the correlation of the states of the units it connects More formally, if s_i and s_j are the states of units i and j after processing stimulus $n - 1$, then the short-term correlational weight c_{ij} on the connection from i to j is set according to

$$c_{ij}^{[n]} = \lambda s_i' s_j' + (1 - \lambda) c_{ij}^{[n-1]} \qquad (1)$$

where $s_i' = 2s_i - 1$ (scaling each unit state to range beween ± 1) and λ is the exponential weighting proportion (0 5 in our simulations) Both the long-term and short-term weights contribute to how units interact Specifically, in processing the next stimulus n, the summed input $x_j^{(t)}$ to each unit j at iteration t becomes

$$x_j^{(t)} = \sum_i s_i^{(t-1)} \left(w_{ij} + \gamma c_{ij}^{[n]} \right) \qquad (2)$$

where γ balances the contribution of the short-term weights relative to the long-term weights [0 1 in our simulations; cf Eq (3), Appendix A] The short-term weights must be kept relatively small so that the standard weights can solve the task regardless of how the short-term weights are set The states of units are computed from their summed input according to the standard sigmoid function [see Eq (4), Appendix A] Notice that the short-term weights do not change over iterations in processing a stimulus, but change only once the network has settled The effect of the short-term weights is to bias the network toward recently occurring patterns of activity Although our simulations involve backpropagation networks that do not formally minimize an energy function (cf Hopfield, 1982), it may help to think of the short-term weights as temporarily lowering the energy (improving the "goodness") of the minimum corresponding to the previous stimulus

There is independent computational and empirical motivation for introducing short-term weights In the domain of object recognition, the most common use of short-term interactions among units is to temporarily bind together combinations of visual features into a coherent whole (Crick, 1984; von der Malsburg, 1981, 1988; von der Malsburg & Schneider, 1986) The recent discovery of synchronized oscillations in the responses of visual cortical cells to disjoint moving contours of a single object (Eckhorn, Bauer, Jordan, Brosh, Kruse, Munk, & Reitboek, 1988; Engel, König, Gray, & Singer, 1990; Gray, Konig, Engel, & Singer, 1989) has led to the development of a number of models of synchronized neuronal activity for feature binding involving short-term interactions among units (Atiya & Baldi, 1989; Baldi & Meir, 1990; Bush & Douglas, 1991; Eckhorn, Reitboek, Ardnt, & Dicke, 1989; Horn, Sagi, & Usher, 1991; Hummel & Biederman, 1992; Kammen, Koch, & Holmes, 1990; Konig & Shillen, 1991; Lytton & Sejnowski, 1991; Sompolinsky, Golomb, & Kleinfeld, 1989; Sporns, Gally, Reeke, & Edelman, 1989; Wilson & Bower, 1990)

Short-term weights have other interesting computational properties Learning with short-term weights can minimize the interference to old knowledge caused by new learning, and can rapidly recover the old knowledge by canceling out this interference (Hinton and Plaut, 1987) Although the procedure employed by Hinton & Plaut for changing the short-term weights depends on the error on the task rather than directly on the states of units themselves, a bias toward previous interpretations would arise if both short- and long-term weights were updated after every stimulus presentation (McClelland & Rumelhart, 1985) Hinton (personal communication, described in McClelland & Kawamoto, 1986) demonstrated

how to use short-term weights to implement recursion in a network that draws shapes composed of other shapes The long-term weights hold the knowledge about how to draw shapes The short-term weights hold context information about what to draw next once the network is finished with drawing the current shape Thus, the short-term weights function like a "stack" that can reinstate the calling context once a drawing "subroutine" returns Short-term interactions have also been employed for recruitment of units during learning (Feldman, 1982)

Short-term weights are also useful in accounting for empirical phenomena in cognitive psychology The most obvious of these are repetition and semantic priming effects, both in normals (Collins & Quillian, 1969; Mandler, 1980; Meyer & Schvaneveldt, 1976) and amnesics (for a review, see Shimamura, 1986) McClelland and Rumelhart (1985) simulate a range of priming effects with immediate changes directly to the long-term weights rather than to a separate set of short-term weights, although the same results would also hold in the latter case (see also McLaren et al , 1989) Another appropriate domain involves short-term memory and its consolidation into long-term memory (e g , Gardner-Medwin, 1989) Goebel (1990) suggests how to use rapidly changing correlational weights for serial rehearsal in short-term memory (see also Schmidhuber, 1992) Cleeremans and McClelland (1991) show how short-term weights can account for the temporary biases of subjects in learning to respond to structured event sequences This last work is particularly interesting because it involves specific biases toward recently occurring *associations* between stimuli, above and beyond the bias changes for the individual stimuli themselves This suggests that the short-term mechanism involves *weights* between units rather than, or in addition to, simple threshold changes for individual units (cf Morton, 1969)

At a neurophysiological level, it has been known for some time (e g , Kupferman, 1979; Hartzell, 1981) that changes in synaptic efficacy at a single synapse occur at many different time scales A great deal of recent research has elucidated the neurobiological mechanisms of associative learning in the form of long-term potentiation (LTP, Bliss & Lømo, 1973; Lynch, McGaugh, & Weinberger, 1984; Cotman & Monaghan, 1988), also known as long-term enhancement (LTE, McNaughton, 1982; McNaughton & Morris, 1987) LTE differs mechanistically from non-associative forms of short-term potentiation (McNaughton, 1983) and decays with a range of time-constants (Barnes & McNaughton, 1980) The rapidly decaying components of LTE could naturally implement the type of learning carried out by the short-term weights in our network

Thus, there is some independent motivation for extending the computational formalism to include short-term correlational weights as a means of introducing temporal interactions between successive stimuli However, it should be kept in mind that in the current context we are extending the formalism in direct response to the observation of perseverations in optic aphasia, and it is in this sense rather ad hoc For this reason, the simple occurrence of perseverations in the network should be viewed as less interesting than the interactions of these perseverative effects with other aspects of the network's behavior, which are not inherent in the extension of the formalism

A SIMULATION OF IMPAIRED OBJECT NAMING

Following the dyslexia simulation, we develop a network for mapping visual representations of objects onto semantic representations, and compare its behavior under damage with that of optic aphasics We begin by describing the details of the task the network is to perform We then describe the network architecture and the procedure by which it is trained Following this, we compare the behavior of the network under damage with the behavior of optic aphasics in visual object naming tasks

The Task

Forty objects were chosen from four categories of indoor objects: kitchen objects, office objects, furniture, and tools (see Table 3) We first summarize their visual (input) representations, and then their semantic (output) representations The full details of these representations are presented in Appendix B

The input representation for objects was designed to coarsely approximate the kind of visual information that would be available for the purposes of object recognition The representation of each object loosely corresponds to a structural description (Marr & Nishihara, 1978; Palmer, 1977), augmented with information about color, texture, absolute size, and more general visual characteristics of the object Table 4 lists the type of information represented by each of the 44 visual features

Table 3. The Objects Used in the Simulations

Objects in Each Category			
Kitchen Objects	*Office Objects*	*Furniture*	*Tools*
cup	*pen*	*chair*	*saw*
spoon	*file*	*table*	*nail*
pan	*paper*	*bed*	*plane*
fork	*book*	*sofa*	*ruler*
knife	*disk*	*stool*	*screw*
bowl	*tape*	*rug*	*awl*
can	*stamp*	*radio*	*axe*
plate	*board*	*tele*	*bolt*
dish	*glue*	*divan*	*nut*
glass	*ink*	*desk*	*vice*

Table 4. The Type of Information Represented by Each Visual Feature

	Types of Visual Features	
Features	*Description*	
1–5	Main component	Shape
6–10	Second component	Shape
11–12		Relative position
13–15		Relative size
16–20	Third component	Shape
21–22		Relative position
23–25		Relative size
26–34	General characteristics	
35–37	Color	
38–39	Texture	
40–44	Absolute size	

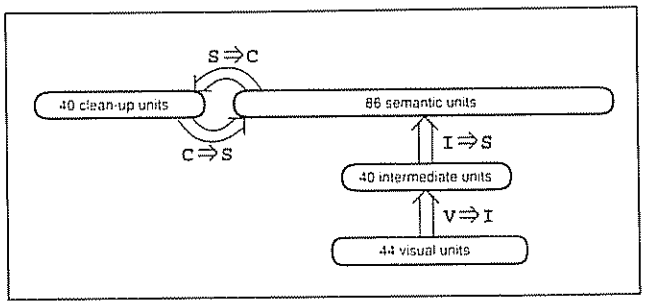

Figure 4. The architecture of the optic aphasia network Notice that the set of connections from the visual (input) units to the intermediate units is labeled V→I rather than G→I

The possible values for each of these types of information are encoded as different patterns of activity over the designated feature groups (see Appendix B)

Much as in the dyslexia simulation, the semantics of each of the 40 objects is represented in terms of a set of semantic features (see Appendix B) Twenty-eight of the 86 features represent the abstract visual semantics of the object The first 14 of these are identical to the general visual characteristics, color, and texture encoding used in the visual (input) representations The next three are a condensed version of the absolute size encoding, and the remaining 11 summarize the shape of the object Following this, there are features for consistency (2), material of which the the object is made (8), where the object is found (10), its general function (10), specific function (22), and the general actions associated with it (7) We assume that information about more specific actions associated with an object would be given a non-semantic, possibly motoric, representation, in the same way that the semantic representation of an object contains only very general visual information

The Network

The architecture of the network we will use to map visual representations onto semantic ones is derived from the dyslexia network, and is shown in Figure 4[3] It has a direct pathway from 44 "visual" input units through 40 intermediate units to 86 semantic output units, and a clean-up pathway attached to semantics with 40 additional units Only a randomly selected 25% of the possible connections between two layers are present Unlike the dyslexia network, the current network has no direct connections among semantic units—all interactions among these units must take place via the clean-up units The network has a total of 4492 connections In addition to the standard long-term learning weight, each connection has a short-term correlational weight that operates as described previously

The Training Procedure

The network was trained using backpropagation through time (see Appendix A) to activate each of the appropriate semantic units to within 0 2 of its correct value over the last three of eight iterations when presented with the visual representation of each object At the end of processing each object, the short-term weights were modified according to Eq (1) In this way each object was presented in the context of the outcome of the presentation of the previous object Objects were chosen randomly without replacement for presentation during a sweep to ensure that they were all presented equally often and in an unbiased order To solve the task, the network must derive a set of long-term weights that enables it to recognize each object when preceded by each other object To the extent that the unit correlations for one object are unrelated to those for the next, the short-term weights effectively act like noise in the weights, forcing the network to develop stronger semantic attractors with the long-term learning weights

Although the operation of the network is deterministic, the random order of object presentations causes performance to vary somewhat over successive training sweeps However, the network reliably satisfied the training criteria after about 9000 sweeps through all 40 objects No attempt was made to minimize the training time

The Lesioning Procedure

After the network had learned to recognized each object, we subjected each set of connections to lesions of a wide range of severity: 0 05, 0 1, 0 15, 0 2, 0 25, 0 3, 0 4, 0 5, and 0 7 The severity of lesion specifies the proportion of existing connections that are selected at random and removed from a given set We followed the dyslexia simulation in applying two criteria to the generated semantic activity to decide whether the output of the damaged network constituted a response or an omission

First, a *proximity* criterion ensured that the corrupted semantics was sufficiently close to the correct semantics of some object Specifically, the cosine of the angle (i e , normalized dot product) between the semantic vector produced by the network and the actual semantic vector of some object (in the 86-dimensional space of semantic features) had to be greater than 0 8 Second, a *gap* criterion ensured that no other object matched nearly as well Specifically, the proximity to the generated semantics of the best matching object had to be at least 0 05 larger than the proximity to any other object If either of these criteria failed, the output was interpreted as an omission; otherwise the best matching object was taken as the response, which either could be the correct object or an error We have not implemented an output system that would map the semantics of objects onto their pronunciations Plaut and Shallice (1991a) discuss the difficulties in developing such a system, and demonstrate that criteria and output systems produce similar effects in the domain of word reading In the current simulations we must resort to applying much more computationally demanding criteria directly to semantics Specifically, the expected proximity of random vectors decreases with increasing dimensionality, so these criteria are somewhat more stringent in the current context (with 86 semantic features) than when applied to the semantics of words (with 68 features) While this will lower the overall rate of explicit responses (correct or error), it should not significantly bias the distributions of error types (see Hinton & Shallice, 1991) We will explicitly verify this later

We would like to measure the performance of the damaged network on each object as stimulus when preceded by every other object We will refer to the preceding object as the "prime " One possible procedure for gathering data is to administer a particular lesion, and then measure performance on all of the objects after using each of them in turn as the prime (i e , setting the short-term weights based on the unit correlations when the prime object is presented to the damaged network) However, this procedure has the drawback that the pattern of errors will be quite similar for each different prime—the tendency for a particular lesion to cause particular errors may dominate any perseverative effects The alternative procedure that we adopt is to administer a different lesion for each prime object This means that data is gathered over 800 instances of a particular type of lesion (40 primes × 20 lesions per prime) rather than just 20 In this way, the effects due to a particular type of lesion are better sampled, while still enabling perseverative effects to emerge Although the first procedure is more analogous to the testing situation for an individual patient, the latter should produce results that better reflect the extent to which lesions to the network *in general* produce the optic aphasic error pattern

For each lesion, we randomly selected and removed the appropriate proportion of connections, presented the prime with the short-term weights set to zero, and set the short-term weights on the basis of the resulting unit activities We then presented each object in turn (with the short-term weights fixed) The network responded correctly if the generated semantics satisfied the proximity and gap criteria for the semantics of the presented object It made an error if the criteria were met for some other object, and it produced an omission if the generated semantics failed either of the criteria

RESULTS

Correct Performance

Figure 5 presents the correct performance of the network after lesions to each set of connections, as a function of lesion severity Overall, lesions to the direct pathway (*visual-to-intermediate,* V→I, and *intermediate-to-semantic,* I→S) are more debilitating than lesions to the clean-up pathway (*semantic-to-cleanup,* S→C, and *cleanup-to-semantic,* C→S) Compared with the correct rates for *disconnection* lesions of the dyslexia network (see Hinton & Shallice, 1991, p 86), the optic aphasia network is much less sensitive to I→S lesions and much more sensitive to S→C lesions This contrast can be understood by recognizing that the network finds it quite difficult to correctly recognize objects when the short-term weights provide a bias toward previous objects It relies heavily on the clean-up pathway to overcome this bias, and is therefore more sensitive to lesions to this pathway

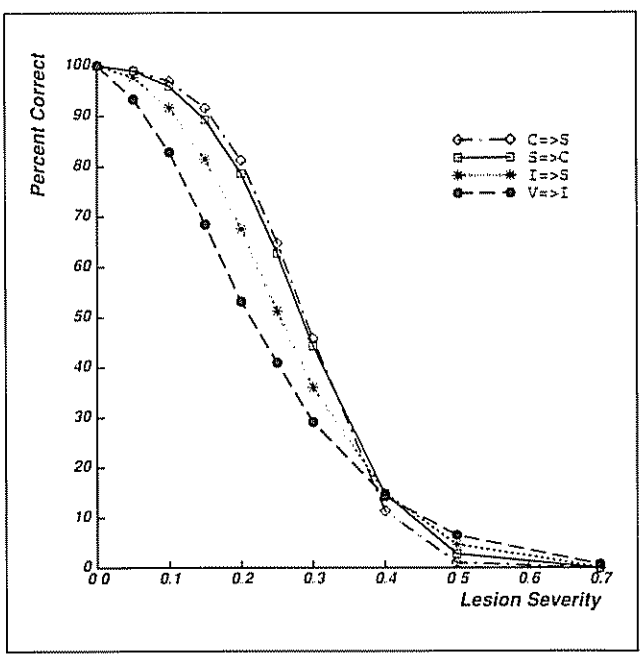

Figure 5. Overall correct performance of the optic aphasia network after lesions to each main set of connections, as a function of lesion severity

Error Classification

When the network is incorrect, it either fails to respond (omission) or produces as a response an object other than the one presented (error) To classify explicit errors, we need a definition of when the stimulus and response are visually and/or semantically related We could use category membership for semantic relatedness as Hinton and Shallice did, but there is no visual similarity measure for objects analogous to letter overlap A more uniform approach is to use the proximity (i e , normalized dot-product) of the representations of the stimulus and response as a direct measure of their visual and semantic similarity We can then apply criteria to these proximity values to determine if two objects are sufficiently similar to be related for the purposes of error classification Accordingly, two objects will be considered visually related if the proximity of their visual representations exceeds 0 6, and semantically related if their proximity of their semantic representations exceeds 0 55 The exact values of these criteria are somewhat arbitrary because we will evaluate the proportions of different error types produced by the network relative to their chance rates We chose these values so that the chance rates of visual and semantic relatedness across all possible object pairs are about equal and near 10% Specifically, these criteria yield the following chance rates of errors: visual v = 0 0949, semantic s = 0 0885, mixed visual-and-semantic m = 0 0615, and other o = 0 7551 Notice that vs = 0 00840 is much less than m If visual and semantic similarity were unrelated, these two values should be about equal, as they are for the Hinton and Shallice word set (Plaut & Shallice, 1991a; 1993)

A more direct test of the relationship between visual and semantic relatedness is the correlation, over all pairs of objects, of visual and semantic proximity In fact, there is a highly significant correlation between visual proximity and semantic proximity for objects [0 52 ignoring diagonal terms, $t(1558) = 23\ 7, p < 0\ 001$] In contrast, the correlation for the word set shows a slight *negative* trend [$-0\ 04$, $t(1558) = 1\ 56, p = 0\ 12$] Thus, a major difference between our definition of object recognition as compared with word recognition is that there is significant structure in the mapping of visual input to semantics for objects but not for words That is, unlike words, objects with similar appearances have similar meanings This will prove important in explaining the rarity of visual errors in optic aphasic object naming compared with visual errors in deep dyslexic reading

Horizontal Errors

We are concerned with two types of effects in the errors produced by the network under damage, roughly corresponding to Lhermitte and Beauvois' (1973) distinction between "horizontal" and "vertical" errors Horizontal errors refer to the standard relation between the stimulus and response For these we will use the definitions of visual and semantic relatedness described above, and classify errors as visual, semantic, mixed visual-and-semantic, and "other" in the same way as Hinton and Shallice (1991) We address vertical (perseverative) errors in the following section

Figure 6 presents the rates of each type of horizontal error after lesions resulting in correct performance between 20 and 80%, as well as the distribution of types in pairs of different stimuli and responses chosen randomly from the set of objects The total error rates are fairly low, ranging from 3 1% for V→I lesions to 0 1% for C→S lesions The very low rates of explicit errors after clean-up lesions indicate that relatively intact attractors are necessary to clean-up the corrupted semantics into those of a related object—with damaged attractors the network either names the object correctly (as was shown in Fig 5) or simply fails to respond Considering the distribution of error types, first notice that proportions of "other" errors—responses unrelated to the stimulus—are very low relative to their chance proportion Visual and semantic similarity have a strong influence on the behavior of the damaged network The clearest effect of this is a strong bias toward mixed visual-and-semantic errors Although their chance rate is only about 6%, they constitute over 68% of the errors produced by the network There is also a strong bias toward semantic errors as compared to visual errors Overall, the ratio of semantic errors to visual errors is 5 7 times the chance ratio This ratio increases as the location of damage moves closer to semantics For V→I the ratio is 4 5 times the chance value, while lesions to the clean-up pathway (S→C and

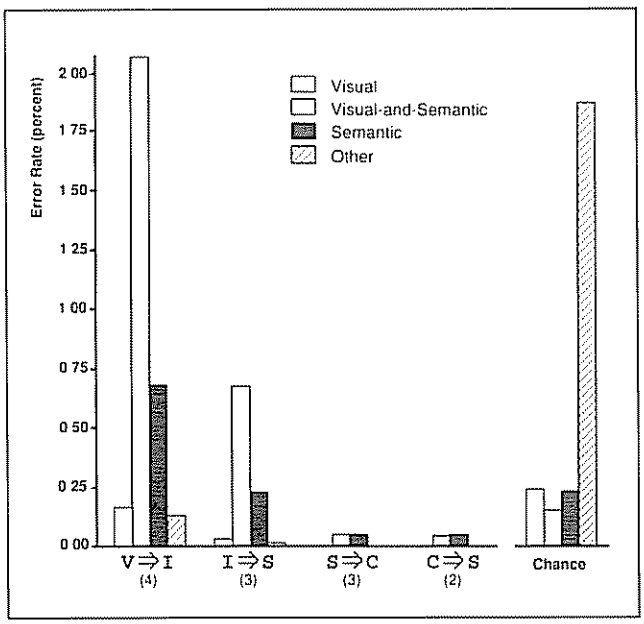

Figure 6. The distribution of error types for lesions to the optic aphasia network producing correct performance between 20 and 80%, as well as the distribution of types for error responses chosen randomly from the set of objects

C→S) produce virtually no visual errors In fact, almost 92% of the errors produced by the network are semantically related to the stimulus, compared with a chance value of 15%

Why, then, does the network show such a strong bias toward semantic and mixed visual-and-semantic errors relative to visual errors? Another way to phrase this question is, why is the bias toward semantic vs visual similarity in errors so much stronger in the optic aphasia network than in the dyslexia network? In the dyslexia simulations, the ratio of visual errors to semantic errors is roughly equal to that of their chance rates (see, e g, Plaut & Shallice, 1991a, p 82) It is unlikely that the difference relates directly to the presence of short-term weights in the optic aphasia network A bias in errors toward the previous object would match the chance distribution of error types, since all combinations of primes and stimuli were tested Rather, it is more likely to relate to a difference in the nature of the tasks of object recognition and word recognition Earlier we argued that the relationship between visual and semantic representations is more structured for objects than it is for words, both in general and in how we have defined the tasks That is, two objects with similar visual forms are more likely to have similar meanings and functions than are words that share letters In a sense this follows the distinction made by Gibson (1979) that shapes have particular "affordances"—they allow for certain types of manipulations and actions independent of specific knowledge of their identity (see also Riddoch & Humphreys, 1987)

We explain word reading errors in deep dyslexia, and object naming errors in optic aphasia, in terms of the same computational principles: damage to an attractor network causes the initial pattern of activity generated by a stimulus to be "captured" by the attractor for a related stimulus (see Fig 3) In mapping orthography to semantics, there is strong pressure to position and shape the attractors so as to separate the initial semantic activity for visually similar stimuli into their quite distinct, final semantics In fact, for many words (e g, BUG), visual similarity with words in other categories (e g, MUG, BOG) is often much greater than with any of the words in the same category (e g, DOG, PIG) Because visually similar words tend to produce similar initial patterns of semantic activity, the network must learn to position their attractor basins to pull apart these similar patterns into quite different final patterns As a result, there are large areas within semantic space where the attractor basins for purely visually related words adjoin, providing ample opportunity for visual errors (see also the analysis in Appendix A of Hinton & Shallice, 1991) In contrast, as discussed above, in object naming it is less common that visually similar objects need to be separated into completely different semantics In addition, even for visually similar, semantically distinct pairs of objects (e g, *fork* and *awl*), there are typically other objects within each category that are just as visually similar (e g, *spoon* and *nail*, respectively) When the initial semantics for fork is corrupted by damage, the additional bias of semantic similarity makes the mixed error *fork* → "spoon" much more likely than the visual error *fork* → "awl" Thus, potential visual errors are often preempted by semantic or mixed visual-and-semantic errors

In fact, the optic aphasia network is more likely to produce semantic errors relative to visual errors than are the deep dyslexia networks, even when the chance possibilities for visual and semantic errors are made approximately equal (compare, for instance, Fig 6 with Fig 5 6 of Plaut & Shallice, 1991a) This seems likely to be due to another difference between the networks The optic aphasia network has to constantly overcome the disturbing influence of the perseverative effects of the preceding stimulus This requires it to build more powerful attractors Since stronger attractors produce more accurate semantic representations, putative errors are more likely to satisfy the gap criterion Semantic errors are particularly sensitive to the gap criterion because they differ from semantically related alternatives (including the correct response) on only a few features Thus, by helping to satisfy the gap criterion, stronger attractors tend to increase the rate of semantic errors relative to visual errors[4]

For these two reasons, then, the optic aphasia network is much more prone to produce semantic errors compared with visual errors than are the deep dyslexia networks This difference corresponds to that obtained between the two neurologically based error patterns that are being considered

Vertical Errors

One of the more interesting aspects of the naming errors of optic aphasics is that they are biased by the responses given to previously presented objects These perseverative errors, termed "vertical" by Lhermitte and Beauvois (1973), are most frequently identical to previous responses, but can also be semantically related Mixed visual-and-semantic perseverative errors occur as well, but purely visual perseverations have not been documented

In the preceding section we analyzed error responses based only on the relationship between stimulus and response—so-called "horizontal" errors Classifying vertical errors is more complicated as they involve the relationship of both the stimulus and response with previous objects For simplicity we will confine ourselves to considering the effects of only the immediately preceding object, which we call the "prime" This simplification also applies to most of the errors produced by patients We will also consider only *semantic* relatedness between the prime and the stimulus and/or response, as the rate of purely visual relatedness is so low Finally, if the damaged network misnames the previous object, we

will consider the *named* object to be the "prime" for the purposes of comparison with the stimulus and response

Table 5 presents the possible relationships between stimulus, response, and prime, and their classification into error types Each type is labeled in two parts The first part (before the "+") refers to the nature of the stimulus–response (horizontal) error:

S Stimulus and response are semantically related (includes semantic and mixed visual-and-semantic errors)

O Stimulus and response are not semantically related (includes visual and "other" errors)

The second part of the label refers to the nature of the (vertical) perseveration, depending on the relationship between the prime and the stimulus and/or response:

P The response is identical to the prime ("perseveration")

$\overline{P}$ The stimulus is identical to the prime but the response is not In this case the prime-response relationship is *contrary* to an item perseveration

C The response is semantically related ("coordinate") but not identical to the prime

$\overline{C}$ The stimulus is semantically related to the prime but the response is not Thus, the response goes against a semantic perseveration

U The response and prime are unrelated

The easiest way to understand these types is in terms of the different influences that contribute to the error An S+U error [e g ("spoon") *desk* → "chair"] is a standard semantic error with no perseverative influence, while an O+P error [e g ("spoon") *desk* → "spoon"] is a repeated response unrelated to the stimulus In an S+C error [e g ("spoon") *fork* → "knife"], both horizontal (stimulus–response) and vertical (prime–response) semantic similarity contribute to the error. In an error involving the perseverative relation $\overline{P}$ or $\overline{C}$ [e g ("spoon") *spoon* → "desk"], the stimulus is consistent with the perseverative effect but the response is not—thus the error is contrary to the perseverative influence

As mentioned in the Introduction (see Table 2), of the types of errors of JF described by Lhermitte and Beauvois (1973), most are horizontal semantic errors (S+U) but both O+P, S+P, and S+C occurred reasonably often as well

Figure 7 shows the distribution of these perseverative error types produced by lesions to the optic aphasia network, averaged over all lesion locations and severities producing correct performance between 20 and 80% The figure also shows the rates of each error type that would be expected if the preceding object had had no influence on errors The predominance of horizontal semantic errors is clear in the figure However, since our current concern is with perseverative influences, we will first consider the remaining ("O") errors These errors provide the clearest picture of perseverative influences because there is no confounding bias from semantic relatedness of the stimulus and response

Pure perseverations (O+P), in which the previous response is repeated even though it bears no relation to either the current stimulus or response, make up 3 73% of the network's errors The overall proportion of error responses that are not semantically related to the stimulus is 8 26% If there were no perseverative influence on errors, only 1 in 40 of these responses would be identical to the prime, and so only 0 21% of all errors would be O+P by chance Thus, the observed rate of response perseverations is about 18 times the chance rate Another indication of the strong perseverative influence is that it is extremely rare (0 02% of errors) for the prime itself to produce an unrelated response when presented as the stimulus (O+$\overline{P}$), even though this type of error is just as likely by chance as O+P Thus, the prime is exerting a strong bias on the nature of the response independent of any relationship with the stimulus

Table 5. The Possible Types of Errors Based on Semantic Relatedness between Stimulus, Response, and Prime

	Relationship			*Example*	
Type	*Stimulus–Response*	*Prime–Stimulus*	*Prime–Response*	*Prime*	*Stimulus → Response*
S+P	Semantic		Identical	"spoon"	*fork* → "spoon"
S+$\overline{P}$	Semantic	Identical	Semantic	"spoon"	*spoon* → "fork"
S+C	Semantic		Semantic	"spoon"	*fork* → "knife"
S+U	Semantic		None	"spoon"	*desk* → "chair"
O+P	None		Identical	"spoon"	*desk* → "spoon"
O+$\overline{P}$	None	Identical	None	"spoon"	*spoon* → "desk"
O+C	None		Semantic	"spoon"	*desk* → "fork"
O+$\overline{C}$	None	Semantic	None	"spoon"	*fork* → "desk"
O+U	None		None	"spoon"	*desk* → "nail"

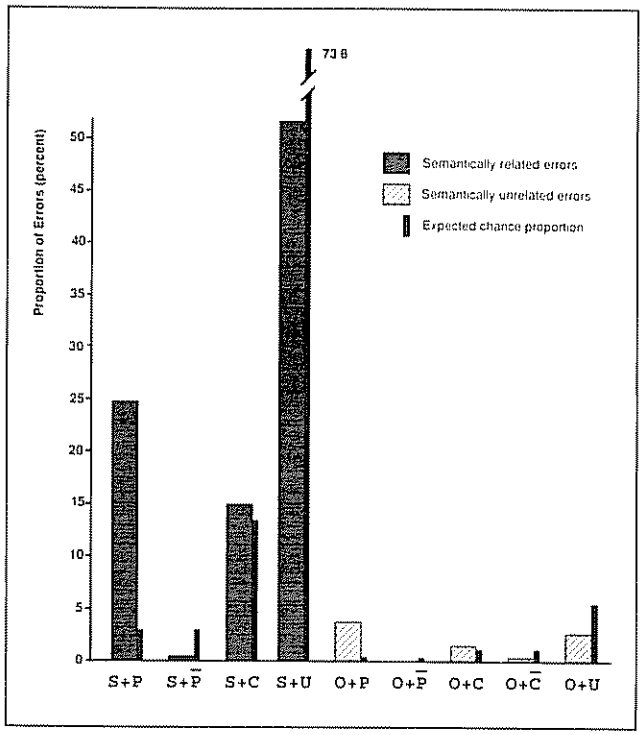

Figure 7. Distribution of perseverative error types, averaged over all lesion locations and severities producing correct performance between 20 and 80%. The black bars indicate the expected rate of each type if there were no perseverative influence.

The network also produces some semantic perseverations (O+C), in which, rather than the prime itself, an object that is semantically related to prime is given as a response unrelated to the stimulus [e.g. ("spoon") *desk* → "fork"; 2.75% of errors]. In contrast, the network is much less likely to produce an unrelated response when the *stimulus* is semantically related to the prime [O+C̄, e.g. ("spoon") *fork* → "desk"; 0.24% of errors]. Both these errors and completely unrelated errors (O+U) occur at rates far below chance. Thus, the prime biases the network toward responses that are semantically related to it—this increases O+C errors and decreases O+C̄ and O+U errors.

A similar pattern of results holds among errors in which the stimulus and response are semantically related. Among these, by far the most common is the conventional horizontal semantic error with no perseverative relationship (S+U; 51.6% of errors), although these fall well below their chance rate. Among the errors for which there is a perseverative influence, the most common are response perseverations [S+P, e.g. ("spoon") *fork* → "spoon"; 24.8% of errors], occurring about 10 times above the chance rate. Semantic perseverations (S+C) are somewhat less common (15.0%) and are only slightly above the chance rate. Thus, the prime induces a strong bias toward an identical response to the next object rather than simply one in the same category. Also notice that it is very rare (0.38%) for the prime to produce another object in the category rather than itself (S+P̄) when presented as the stimulus. Thus, even within a category, the prime biases responses toward itself compared with other objects in the category. This bias toward S+P errors is also seen in the error responses of JF.

Effects of Type of Response to the Preceding Object

For the purposes of categorizing perseverative errors, we have define the "prime" to be the response given to the preceding object (correct or an error), or the object whose semantics are nearest those generated by the network in the case where presentation of the preceding object resulted in an omission. While we have grouped these conditions together in the analysis presented above, it would seem likely that the influence that the prime has on the naming of subsequent objects would vary considerably with how well the network responded to the prime itself. To investigate this possibility, we separated errors based on how the damaged network responded to the prime.

Figure 8 presents the same data on the distributions of perseverative error types, now separated by whether the network named the preceding object correctly, made an error, or failed to respond. Consider the balance of responses that are identical to the prime (P) vs. those that are unrelated (U), both for semantic errors (S) and other errors (O). This provides a rough measure of the "strength" of perseverative influences. First notice that the proportions of all errors that are response perseverations (S+P and O+P) are much lower when the prime is an omission than when it is an explicit response. Conversely, when the prime produces no response as a stimulus, the proportion of errors showing no perseverative influence (S+U and O+U) are much higher and close to their chance proportions (relative to S and O responses, respectively). In fact, the proportion of all error responses that are identical to the prime is 50.3% for primes producing explicit response vs. only 3.1% for primes producing omissions. Thus, when an object generates no response, it also has far less influence on the naming of subsequent objects than when it generates a correct or error response. This makes sense given that omissions occur when the semantics generated by the network do not match the nearest object very well. When the short-term weights are set on the basis of poorer semantics, they provide a weaker bias toward the prime than when set by more accurate semantics. For this reason, the combined data reported in Figure 7 significantly underestimate the perseverative influence from previous objects that evoke an explicit response.

Even among explicit responses, there are some interesting differences between correctly vs. incorrectly named primes. When the preceding object is named correctly, the proportion of errors involving a response perseveration is about equal to that involving no persev-

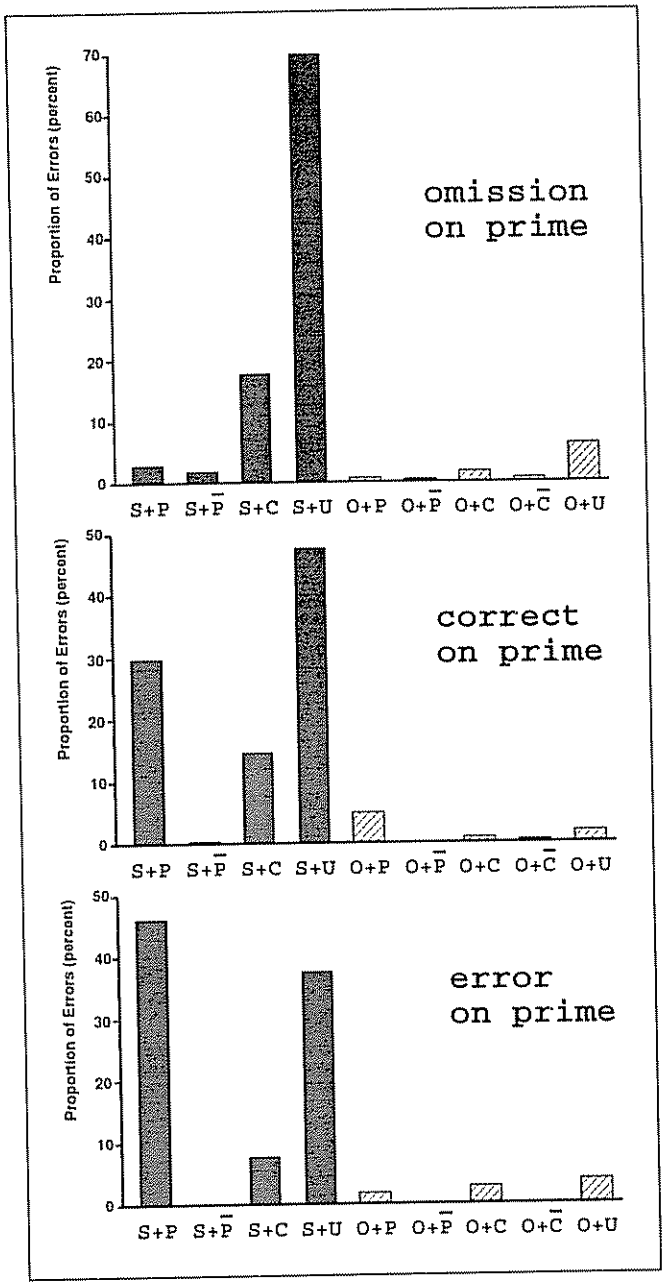

Figure 8. Distributions of perseverative error types following prime presentations resulting in omissions, correct responses, and errors

erative relation In contrast, when the preceding object is named incorrectly, the next object is much more likely to elicit the same response The proportion of errors that are response perseverations is 48 1% for primes producing errors vs 34 4% for primes named correctly In essence, the attractor for the incorrect response has become abnormally strong as a result of damage, producing the error to the preceding object When the short-term weights are set on the basis of this object's semantics, there is even stronger pressure for other objects to succumb to the same attractor

The perseverative responses of optic aphasics can come from correctly named objects, incorrectly named objects, and have even been reported on objects generating no response (Gil et al , 1985; Lhermitte & Beauvois, 1973) The proportions of error types that follow each of these conditions has not been analyzed in detail, so it is difficult to compare the network's behavior with that of patients in more than a qualitative manner However, quite a common effect clinically is a chain of semantically related responses, which could be explained by the existence of an abnormally deep attractor

Effects of Lesion Location

The data from different lesion locations are averaged together in the results presented above However, the distribution of perseverative error types differs significantly as a function of the location of damage in the network Figure 9 presents the distributions of these error types separately for each lesion location The pattern for V→I lesions is most similar to that for the entire network because the largest proportion of errors (73 2%) occur after these lesions There is an interesting progression in the error pattern as the lesion location moves closer to semantics For V→I lesions, most semantic errors show no perseverative relationship, while other errors are about balanced between O+P and O+U For these lesions, the proportion of error responses that are identical to the prime is 27 7% I→S lesions show a slightly higher proportion of perseverative responses (30 2%) Clean-up lesions produce virtually no nonsemantic errors, and those that occur are either item or semantic perseverations with the prime Overall, compared with lesion to the direct pathway, clean-up lesions produce a higher proportion of perseverations (33 3% after C→S lesions and 42 3% after S→C lesions) Thus, the strength of the perseverative influence increases as lesions move closer to semantics However, the proportion of semantic perseverations (S+C and O+C) is somewhat less affected by lesion location

Why should lesions near or within semantics produce a stronger bias toward response perseverations than lesions closer to the input? Clean-up lesions corrupt the semantic attractors for objects, resulting in far fewer overall naming errors than do lesions to the direct pathway (see Fig 6) After clean-up lesions, the prime is named correctly on 71 8% of the trials.[5] When this occurs, the short-term weights within the clean-up pathway are set in a way that magnifies the clean-up influences that generated the semantics of that particular object This additional bias has more influence after clean-up lesions compared with direct-pathway lesions because the normal clean-up influences are diminished after the former but not the latter The bias toward the semantics of the preceding object can dominate the weakened clean-up for the correct semantics of the stimulus, causing the network to more frequently produce a response perseveration Even when the semantics generated by the network in response to the preceding object do not satisfy

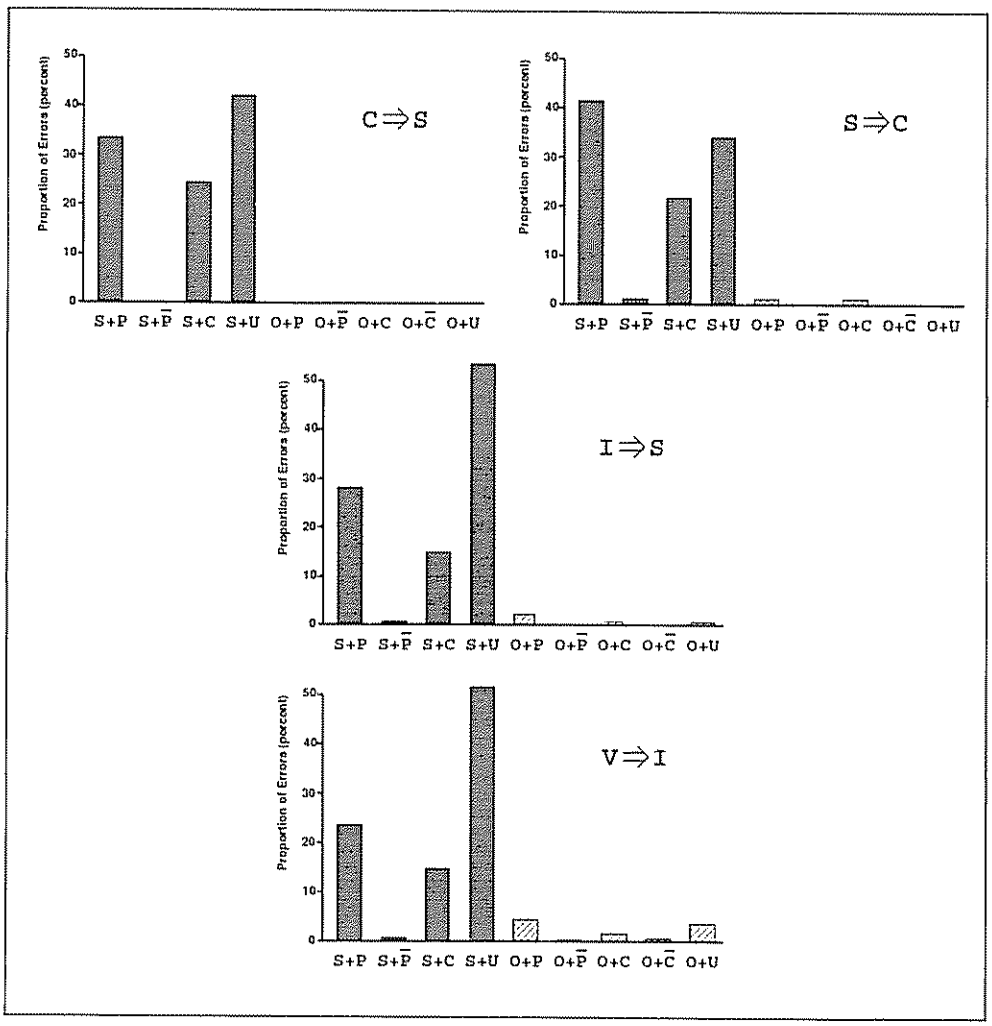

Figure 9. Distributions of perseverative error types for each lesion location

the response criteria, they still evoke short-term weights that bias the network on the next trial toward the semantics of that object as compared with others

Effects of Severity of the Impairment

All of the results we have presented thus far have been averaged across lesions producing overall correct performance between 20 and 80% Our motivation for considering only this range of performance is that quite mild or severe impairments often produce an atypical distribution of error types (see Plaut & Shallice, 1991a) Also, the levels of correct performance of most of the patients we are considering fall into this range However, since patients vary significantly in their overall levels of performance, it is important to know how the behavior of the network varies over a range of severities of impairment as well

There are two general trends in the distribution of errors as performance deteriorates with increasing lesion severity First, the proportion of error responses that are unrelated to the stimulus gradually increases Only 2 7% of all errors are unrelated when correct performance is above 80%, while over 22 2% are unrelated when performance is below 20% Second, as performance deteriorates, the proportion of errors that are response perseverations gradually decreases S+P and O+P errors account for 47 3% of all errors when correct performance is above 80%, but only 20 9% when performance is below 20% Thus, as the severity of the impairment increases, both the stimulus and the preceding object have diminished influence on the responses generated by the network Conversely, networks with only slight impairments are particularly prone to producing responses that are related to *both* the stimulus and prime—58 5% of all errors are either S+P or S+C when correct performance is above 80%

Effects of Response Criteria

To verify that our results do not depend on the particular values of the response criteria used, we reevaluated the behavior of the network under the identical set of lesions, using less stringent criteria for explicit responses Specifically, we decreased the proximity criterion from 0 8 to 0 75, and decreased the gap criterion from 0 05 to

0 025 These relaxed criteria increase the overall proportion of word presentations producing explicit responses from 51 2 to 61 1% Although both the rates of correct and error responses increase, errors account for a higher proportion of explicit responses when the criteria are more lenient (6 8 vs 3 7%)

If we consider lesions producing correct performance between 20 and 80%, the network produces a much higher overall error rate when using less stringent criteria (4 0 vs 1 3%) However, the distribution of error types is quite similar The network continues to show a strong bias towards mixed visual-and-semantic errors (59 4% of all errors) and many more semantic than visual errors, with the ratio being four times the chance ratio These numbers are somewhat lower than those for the original criteria (68 2% and 5 1) because the less stringent criteria produce a higher proportion of "other" errors (8 9 vs 3 6% of all errors, respectively)

The basic perseverative effects also remain when using relaxed response criteria The rates of S+P, S+C, O+P, and O+C errors are all still well above the rates expected if the prime had no effect on the response, while S+$\overline{P}$, O+$\overline{C}$, and O+$\overline{P}$ errors occur at rates well below their expected chance rates Overall, 19 5% of all errors are item perseverations when using the less stringent criteria, compared with 28 5% with the original criteria In fact, completely unrelated errors are more common with less stringent criteria (6 6 vs 2 7% of all errors) Thus, while less stringent criteria allow poorer semantic activity patterns to qualify as responses, thereby reducing the overall influence of the stimulus and prime on the response in the averaged data, the same pattern of horizontal and vertical effects in errors is present

DISCUSSION

The current simulation demonstrates that the computational approach used in our earlier work on deep dyslexia can be extended to the analysis of the error pattern of optic aphasia In particular, given our earlier work, one could expect that the errors that occur when the optic aphasia network is lesioned would reflect the similarity metrics of the input and the output representations, both separately (i e , in semantic errors and visual errors) and together (in mixed visual-and-semantic errors) And in fact, semantic errors and visual errors occur at well above chance levels, and mixed visual-and-semantic errors occur even more frequently

In both the dyslexia and optic aphasia simulations, a small set of general connectionist principles—relating to distributed representations, gradient descent learning, and attractors—is sufficient to account for complex empirical phenomena concerning impaired visual comprehension This strongly suggests that these principles shed light on the nature of the computations underlying the semantic processing of visual information

The major extension in the present simulations is the incorporation of short-term correlational weights In addition to having other interesting computational and empirical consequences, these weights bias the network toward reproducing patterns of activity that have occurred on recently preceding trials As a result, the effects of lesions to the network have an additional dimension Responses that have occurred on the preceding trial now occur as errors at well above chance rates, especially when the stimulus is semantically related to the preceding response In addition, less frequently but still above chance, a coordinate of the preceding item occurs as an error response If one considers the optic aphasic error pattern described in the Introduction, both of these types of error have been extensively documented They correspond to the two types of "vertical" error discussed by Lhermitte and Beauvois (1973) Frequent semantic errors are also a standard part of the optic aphasic error pattern, and a high frequency of mixed errors have been noted by Riddoch and Humphreys (1987) in their patient JB They also noted that the patient had particular difficulty in discrimination between objects that were both visually and semantically similar, which they also interpret in terms of cascaded processes

Three main empirical issues remain with respect to the syndrome: whether the simulation reproduces all aspects of the error pattern, how the two exceptional patients referred to in the Introduction can be explained, and what account can be given of other aspects of the syndrome However, before turning to these questions, it is important to clarify how the network architecture and lesioning procedure relate to the underlying neuroanatomy of object recognition and the typical neuropathology in optic aphasia

Relation to Neuroanatomy and Neuropathology

The main emphasis of our work is to elucidate the general computational principles that underlie the ability of damaged connectionist networks to reproduce neuropsychological phenomena However, for the behavior of a damaged network to be relevant to the behavior of neurological patients, there must be some relationship between the structures that are damaged in the network and the brain structures that are damaged in the patients The general question of the relation between various types of connectionist networks on the one hand, and various aspects of neurobiology on the other, is far beyond the scope of this paper (for discussion, see e g , Allport, 1985; Churchland & Sejnowski, 1988; Crick, 1989; Crick & Asanuma, 1986; Nadel, Cooper, Culicover, & Harnish, 1989; Smolensky, 1986, 1988) Nonetheless, we can make some tentative suggestions as to how groups of units in the current network might relate to neuroanatomy

The visual input units in our network represent high-level visual information in the form of structural descrip-

tions of objects Substantial neurophysiological and neuroanatomical evidence, from both human and animal work, suggests that the highest-level presemantic visual representations are found in inferior temporal cortex (see Plaut & Farah, 1990, for a review) Based on an extensive review of reported cases of associative agnosia, Farah (1990, 1991) proposes bilateral involvement in humans, with each hemisphere specializing in different types of information However, Warrington and Taylor (1973, 1978) consider right hemisphere systems to be particularly important, at least for complex stimuli

Not surprisingly, localizing semantics is more problematic Of course, one possibility (see e g , Allport, 1985) is that semantics cannot be localized to any particular brain location, but is distributed across a number of cortical areas that are specialized for different modalities to different extents In fact, most generalized dementias, such as Alzheimer's disease, involve diffuse cortical damage (for an overview, see Heilman & Valenstein, 1985) However there is increasing evidence that left temporal structures are particularly important for semantic processing A series of patients with unilateral left hemisphere lesions were given a set of verbal tests by Coughlan and Warrington (1978) The three tests that had low semantic loading were performed roughly equally well by patients whose lesions involved the temporal lobe and by those whose lesions spared the temporal lobe However the four tests that had a strong semantic component were performed significantly worse by the temporal group More recently, there have been a series of studies of individual progressive aphasic patients using CT, MRI, and PET Where the language problem was well restricted to semantic processing, focal atrophy of the left temporal lobe has been the standard finding (e g , Poeck & Luzzatti, 1988, cases 2 and 3; Tyrrell, Warrington, Frackowiak, & Rossor, 1990, cases 2 and 5; Hodges, Patterson, Oxbury, & Funnell, 1992)

The one finding which appears discrepant with a key role for the left temporal lobe in semantic processing is the suggestion from work using PET imaging that semantic processing involves the left dorsolateral prefrontal cortex (Petersen, Fox, Mintun, & Raichle, 1988) However, the task employed required subjects to generate a use for a noun (e g , CAKE → "eat") Frith, Friston, Liddle, and Frackowaik (1991) found that generation tasks that lack a semantic component also involve the same part of the frontal lobe They argue that it is this component of the task used by Petersen and colleagues, and not the semantic aspect, that leads to left dorsolateral frontal activation This certainly fits with the classical neuropsychological view of word fluency tasks (see Milner, 1964)

One of the main strengths of connectionist neuropsychology is that brain lesions have a natural, fairly atheoretical interpretation in terms of removal of units and/or connections (but see Small, 1991, for alternative interpretations) Thus, to the extent that we can associate groups of units in the model with brain areas, we would interpret lesions to the model as corresponding to partial or complete lesions of the corresponding brain structures On this basis, our simulation of the optic aphasic error pattern after lesions that impair semantic processing of visual information fits well with the neurological evidence that most known optic aphasic patients have unilateral left posterior damage, and that this region is particularly important for semantic processing

With this relationship between network structure and brain structure in mind, we return to considering the extent to which the simulation reproduces all aspects of the optic aphasia error pattern To this end, it is useful to compare optic aphasia and deep dyslexia, and their simulations, with respect to the relationships between their respective input and output representations

A Comparison of Optic Aphasia and Deep Dyslexia

The two main differences in the pattern of performance of deep dyslexics and optic aphasics that are relevant relate to the relative frequency of purely visual errors and perseverations Errors that are visually but not semantically related can constitute a fairly high proportion of all errors in some deep dyslexics (e g , 51% for PS, Shallice & Coughlan, 1980) and yet are quite rare in optic aphasics (Riddoch & Humphreys, 1987; Gil et al , 1985; Lhermitte & Beauvois, 1973) One possible contribution to this difference is that the definition of visual similarity may be more stringent for objects than for words However, with presumably the same criteria for visual similarity, some other visual agnosics make a high proportion of visual errors in naming objects (e g , 46% for FZ; Levine, 1978; also see Larrabee, Levin, Huff, Kay, & Guinto, 1985) Thus, the rarity of visual errors by optic aphasics cannot be completely explained by a criterion difference Our explanation is that the relative difference in visual errors in reading vs object naming is due to the different amount of structure in the two tasks and the different overall strength of the attractors As demonstrated in the simulations, the greater similarity between the visual and semantic representations of objects diminishes the influence of purely visual similarity on the layout of attractor basins within semantics, as reflected in the pattern of errors produced under damage The stronger attractors at the semantic level also lead to a greater increase in semantic errors

Is it legitimate to assume there is greater structure in the relationships between the visual and semantic representations of objects compared with words? In our dyslexia simulations we have assumed that the mapping from orthography to semantics is completely unstructured, which is only approximately true However, many aspects of the visual representation of an object are semantically relevant as well Strong similarity of visual form is a characteristic of many semantic categories, par-

ticularly biological ones Even members of functionally defined categories often share visual characteristics because similar shapes are appropriate for similar actions (e g , elongated shapes for pounding, horizontal shapes for sitting/lying, Gibson, 1979) Moreover, size is clearly semantically relevant as are many of the general visual characteristics we used Thus, the assumption that object naming is more structured than word naming seems justified (also see Bobick, 1987; Richards, 1988, for discussion)

The second main difference between reading in deep dyslexia and object naming in optic aphasia relates to perseveration Although perseveration is common after many types of language-related impairments (Albert & Sandson, 1986), it is not particularly prevalent in the reading errors of deep dyslexics In contrast, a relatively high proportion of the naming errors of optic aphasics are related to previously presented objects (e g , 28% for JF, Lhermitte & Beauvois, 1973) We introduced short-term correlational weights to provide a means by which object naming could be influenced by the responses given to previous objects While some amount of independent justification can be given for such weights, we were motivated to include them in the current simulations directly by the observation of perseverative effects in patients, and so they must be viewed as somewhat ad hoc However, the fact that they lead to interesting interactions with other aspects of the operation of the network, such as semantic influences in errors, suggests that their introduction contributes in a significant way to understanding the nature of perseverative influences in optic aphasia The simulation demonstrates that these two differences can account for the relative rarity of purely visual errors in optic aphasia Thus, the whole error pattern can be explained on the present approach

However, a question remains: if there is independent motivation for including short-term weights in mapping between visual and semantic representations in object naming, why did we not include them in networks for mapping orthography to semantics in reading? The simple answer is that the behavior of the patients with impairments in this mapping is better explained without them However, this answer is unsatisfying without an independent explanation for why a temporary bias toward previous patterns of activity is *computationally* appropriate in object recognition but not reading One possible explanation is that, unlike in object recognition, in reading for meaning there is great pressure to recognize successive words as quickly as possible As long as the meaning of each individual word is unrelated to the next, any bias toward the semantics of previous words would induce a kind of "sluggishness" that would impede the network in deriving the correct semantics of the current word More generally, short-term weights are not appropriate in a network for a task in which the speed of separate successive interpretations is critical However, the infinite generative capacity of language has no equivalence in object identification Successively recognized objects—those found together—will tend more to be related than successive words, so a temporary bias in object recognition would be beneficial In addition, the use of short-term interactions for feature binding and segmentation is more critical for natural objects than for words Finally, the pace at which dyslexics are tested is far slower than the normal rate of word recognition, while object recognition is tested at far closer to its "natural" rate

Individual Cases

As mentioned in the Introduction, there are at least two patients who, in terms of their overall pattern of performance, would be classified as optic aphasics or associative agnosics, and yet whose performance presents problems for the current account These are the optic aphasic patient studied by Coslett and Saffran (1989) and the associative agnosic patient, FRA, studied by McCarthy and Warrington (1986)

Coslett and Saffran's patient was virtually entirely unable to name objects, and most of his errors (39/49) bore no semantic or visual relation to the stimulus (e g , *scissors* → "clocks"; *volcano* → "pillar") Seven responses were semantic errors, two were perseverations, and one was a visual error Thus, on the model, he would clearly need to have a severe lesion Yet on a number of object comprehension tests—categorization tasks, functional similarity judgments, and semantic association judgments—he performed well

However, if one examines the patient's performance in more detail, the contrast between spared nonverbal performance and abolished naming performance becomes less striking First, it should be noted that the patient was far from perfect (50% correct) at gesturing appropriately to objects, and the apparent superiority over naming does not necessarily conflict with the model, as we will see in a moment Second, as previously mentioned, in our simulation the proportion of error responses unrelated to the stimulus increases with lesion severity (see also Hinton & Shallice, 1989; Plaut & Shallice, 1991a) For example, V→I(0 5) lesions, which reduce correct performance to 6 5%, produce over three times more errors that are completely unrelated to the stimulus and prime than do lesions producing correct performance between 20 and 80% [6] Of the remaining errors, 89% are semantically related to the stimulus and 4 5% are pure perseverations (O+P) Yet even when an object is not named correctly, on average the generated semantics is closer to the correct semantics than to those of a randomly chosen unrelated object 89% of the time Thus, the residual semantic activity in the damaged network could support quite reasonable performance on nonverbal tests of comprehension even when overt naming is virtually abolished, similar to Coslett and Saffran's patient

The second patient, FRA of McCarthy and Warrington (1986), presents a different problem Not only was FRA unable to name objects reliably, he also had some difficulty with tasks requiring nonverbal accessing of semantic representations For example, he made many errors at matching two examples of the same object that were very visually different (e g , two types of *razor*) Yet tasks that stressed the structural description level were performed well Thus, his deficit would appear to be within the domain of the model Yet his explicit errors in one experiment are described as "semantic approximations" with "no evidence of perseverative responding "

Clearly since FRA was 50% correct at naming pictures, his deficit is not as severe as that of Coslett and Saffran's patient It is puzzling that his performance on nonverbal semantic judgment tasks should be the worse, particularly on category judgment tasks that both patients carried out (although in different forms) Moreover, FRA's pattern of errors was different from that typical of the syndrome He made either omissions or superordinate semantic errors One possibility is that FRA's lesion lay in the semantic system itself, as McCarthy and Warrington (1986) argue It is not possible to assess the effect of such a location of lesion properly on the present model because intact semantic units are required for the calculation of the proximity and gap measures However, the effect would presumably be similar to S→C lesions (see Plaut & Shallice, 1993) As these lesions give low absolute rates of explicit errors, and almost 98% of these have a semantic component, the absolute rate of nonsemantic errors would be swamped by semantically related ones In addition, accessing an incomplete semantic representation would presumably lead to many superordinate semantic errors

However, as FRA performed normally on naming from auditory description, such an account would seem to entail acceptance of some amount of specialization within semantics (see Beauvois, 1982; Shallice, 1987, 1992; Warrington & Shallice, 1984; but also Caramazza et al , 1990; Riddoch et al , 1988, for criticisms) In addressing these wider issues it is useful to consider other aspects of the optic aphasia syndrome

The Overall Pattern of Performance

A major issue remains to be addressed, regarding the relationship between the current simulation and the *preserved* abilities of optic aphasics The impaired visual object naming of optic aphasics is perplexing because their visual *recognition* of objects, as indicated by gesturing or categorization tasks, as well as their naming of objects presented in other modalities, appears relatively intact

The current simulation reproduces the error pattern of optic aphasics in visual object naming by introducing an impairment in deriving semantics from visual input In this sense it follows Riddoch and Humphreys' (1987) claim that optic aphasia is more appropriately considered a "semantic access agnosia" However, the current research simulates neither intact visual recognition nor intact nonvisual naming in the context of impaired visual naming In what sense then is it a simulation of optic aphasia? The honest answer is that it is not one—it is only a simulation of the error pattern of optic aphasics However, such a simulation is interesting only as an explanation of patient behavior if it can plausibly be extended to incorporate the remaining characteristics of the syndrome

In attempting to reconcile the current simulations with the preserved abilities of optic aphasics, we must reemphasize that the visual recognition capabilities of optic aphasics may not be as intact as generally thought The claims of intact recognition in optic aphasia have been based almost entirely on their performance in tasks involving either gesturing or semantic categorization Regarding categorization, in our discussion of Coslett and Saffran's (1989) patient we provided evidence that nonverbal tests of comprehension could be performed quite accurately by the damaged network even when explicit naming is severely impaired (see also Hinton & Shallice, 1991) We approximated these tests by applying a less stringent "best-match" criterion to semantics, so our explanation assumes that the categorization tasks at which optic aphasics succeed require less precise semantics than naming Riddoch and Humphreys provide some evidence for this by showing that their patient JB was significantly impaired at a categorization task that required distinctions *within* a category (as naming must)

Turning to gesturing tasks, in fact adequate gesturing to misnamed objects has been demonstrated in only three cases (Gil et al , 1985; Lhermitte & Beauvois, 1973; Riddoch & Humphreys, 1987)—in two others (Coslett & Saffran, 1989; Larrabee et al , 1985) gestures incorrectly corresponded to the named object, and in an additional two cases (Assal & Regli, 1980; Peña-Casanova & Roig-Rovira, 1985), gesturing and naming visual stimuli were equally impaired To explain relatively preserved gesturing, Riddoch and Humphreys (1987) point out that gesturing is often judged less stringently than naming, and typically requires less precise semantics Their patient JB was 75% correct at gesturing but only 45.5% correct at naming in a task in which objects were selected to have fully discriminable gestures Thus, gesturing shows some impairment, but is still better than naming

However, the same argument is unlikely to account for the relatively preserved gesturing of some optic aphasics Lhermitte and Beauvois' (1973) patient JF never made an incorrect gesture to a set of 100 pictures of objects of which 31 were misnamed If gesturing were based entirely on the same impaired semantics that underlies poor naming, occasional gesturing errors would be predicted Riddoch and Humphreys (1987) propose that correct gesturing may often be based directly on the (nonsemantic) structural description derived by vision

However, as Farah (1990) points out, often quite different gestures are appropriate for visually similar objects (e.g., a *sewing needle* and a *toothpick*), making it unlikely that visual representations alone can support gesturing.

A possible resolution is to suggest that gesturing in optic aphasics may be based on a combination of intact visual structural descriptions and *residual* semantics. The residual semantics could narrow the range of gestures that are appropriate for the shape of the object to those that are consistent with the general semantics of the object, preventing inappropriate gestures. This would be analogous to the claim that residual operation of the phonological route in phonological alexics "edits out" any potential semantic errors that might arise from the impaired operation of the semantic route (Newcombe & Marshall, 1980). However, there is yet another objection that can be raised to this proposal (see Shallice, 1988b). If gesturing is based on the same semantic representation that supports naming, then there are semantic and perseverative errors that would support a different mime from the one produced. On these trials, could the patient really rely on the semantic representation to select appropriately between different gestures that are supported by the structural descriptions?

Perhaps gesturing in optic aphasia can be based on the intact generation of functional portions of an object's semantics (see Farah & McClelland, 1991; Sacchett & Humphreys, 1992; Warrington & Shallice, 1984) even though the generation of other portions of semantics that normally support naming is impaired. Again, this would involve accepting some amount of specialization within semantics. A simulation based on this proposal would involve naming and gesturing to visual and nonvisual stimuli. It would constitute a complete simulation of optic aphasia consistent with the current account of the error pattern in object naming. However, it remains to be developed.

APPENDIX A: BACKPROPAGATION THROUGH TIME

This appendix gives the mathematical details of the "backpropagation through time" learning procedure (Rumelhart et al., 1986; Williams & Peng, 1990).

The Units

Let $x_j^{(t)}$ be the total input of unit j at time t, and let $y_j^{(t)}$ be its output. Then if w_{ij} is the weight on the connection from unit i to unit j, then

$$x_j^{(t)} = \sum_i w_{ij} y_i^{(t-1)} \tag{3}$$

$$y_j^{(t)} = \sigma(x_j^{(t)}) = \frac{1}{1 + e^{-x_j^{(t)}}} \tag{4}$$

The Forward Pass

The network runs for a fixed number of iterations t_{max}. The input is presented to the network by setting $y_i^{(t)}$ for each input unit i and every t as specified by the input. $y_j^{(t)}$ of the remaining units are initialized to some constant value (0.2 in our simulations). Then for $t = 1$ to t_{max}, unit inputs and outputs are calculated according to Eqs. (3) and (4), respectively.

The Error Function

In addition to the states of the input units, the environment specifies the desired states $d_j^{(t)}$ of each output unit j for some times t (typically the last three iterations in our simulations). The error $E^{(t)}$ for time t, called the *cross-entropy* (Hinton, 1989), is defined over output units j to be

$$E^{(t)} = \sum_j [d_j^{(t)} \log(y_j^{(t)}) + (1 - d_j^{(t)}) \log(1 - y_j^{(t)})] \tag{5}$$

where the total error $E = \Sigma_t E^{(t)}$.

The Backward Pass

The backward pass calculates the derivatives of the error with respect to the states and weights in the network. The error derivative of a unit's state has two components: the derivative of the "external" error function (which is 0 for nonoutput units and for iterations without desired states) and the derivative of the error caused by the unit's influence on other units. The error derivatives for weights have two corresponding terms. Specifically, for $t = t_{max}$ to 1, the derivatives of the error at time t with respect to the states and weights of each unit are calculated according to the following equations:

$$\frac{\partial E^{(t)}}{\partial y_j^{(t)}} = \frac{1 - d_j^{(t)}}{1 - y_j^{(t)}} + \frac{d_j^{(t)}}{y_j^{(t)}}$$

$$\frac{\partial E^{(t)}}{\partial y_i^{(t-1)}} = \sum_j \frac{\partial E^{(t)}}{\partial y_j^{(t)}} \frac{\partial y_j^{(t)}}{\partial x_j^{(t)}} \frac{\partial x_j^{(t)}}{\partial y_i^{(t-1)}} = \sum_j \frac{\partial E^{(t)}}{\partial y_j^{(t)}} y_j^{(t)} (1 - y_j^{(t)}) w_{ij}$$

$$\begin{aligned}\frac{\partial E^{(t)}}{\partial w_{ij}} &= \frac{\partial E^{(t)}}{\partial y_j^{(t)}} \frac{\partial y_j^{(t)}}{\partial x_j^{(t)}} \frac{\partial x_j^{(t)}}{\partial w_{ij}} + \frac{\partial E^{(t)}}{\partial y_j^{(t-1)}} \frac{\partial y_j^{(t-1)}}{\partial x_j^{(t-1)}} \frac{\partial x_j^{(t-1)}}{\partial w_{ij}} \\ &= \frac{\partial E^{(t)}}{\partial y_j^{(t)}} y_j^{(t)} (1 - y_j^{(t)}) y_i^{(t-1)} \\ &\quad + \frac{\partial E^{(t)}}{\partial y_j^{(t-1)}} y_j^{(t-1)} (1 - y_j^{(t-1)}) y_i^{(t-2)}\end{aligned}$$

$$\frac{\partial E}{\partial w_{ij}} = \sum_t \frac{\partial E^{(t)}}{\partial w_{ij}} \tag{6}$$

Weight Updating

The procedure defined by Eq. (6) is applied to each example in turn, accumulating error derivatives for the

weights. At this point, each weight in the network is changed according to:

$$\Delta w_{ij}^{[n]} = -\epsilon \left(\frac{\partial E}{\partial w_{ij}} \right)^{[n]} + \alpha \Delta w_{ij}^{[n-1]} \qquad (7)$$

where ϵ determines the overall learning rate (0.001 in our simulations), α is a *momentum* term that causes weight changes to be similar to previous weight changes (0.95 in our simulations), and n is the number of "sweeps" through the examples so far.

Training Criterion

Because the sigmoid unit function asymptotes at 0 and 1, units cannot actually achieve these values in principle. Accordingly, for tasks with binary outputs (as are ours), training is halted when, for every example, the difference between the actual and desired state of each output unit is smaller than some tolerance (0.1 in our simulations).

APPENDIX B: REPRESENTATIONS

This appendix provides details on the input and output representations used in the simulations described in this paper.

Visual Representations of Objects

Visual input for objects is represented over 44 visual features, as summarized in Table 4. The first 25 features are devoted to representing the shape of the object in terms of up to three "components," one of which is designated as the main component. These might be thought of as loosely corresponding to Biederman's (1987) "geons." The shape of each component is encoded over five units, as shown in Table 6. The position and size of the second and third components relative to the main component are described in terms of two and three additional units, respectively (see Table 7). The remaining 19 of the 44 features describe more general visual characteristics of the object, as well as color, texture, and absolute size information (see Table 8).

Table 9 describes the visual representations of each of the 40 objects in terms of the codes listed in the tables for values of each type of information. Figure 10 shows the actual assignment of each of the 44 visual features to each object.

Semantic Representations of Objects

The semantics of each of the 40 objects is represented in terms of a set of semantic features, listed in Table 10. Figure 11 shows the assignment of semantic features to each of the 40 objects.

Table 6. The Encoding Used to Describe the Visual Shape of Each of the Three (Possible) Components of Each Object[a]

Component Shape						
Features					*Code*	*Description*
1	1	1	1	1	cy	Cylinder
1	1	1	1	0	cyh	Cylinder—hollow
1	1	1	0	1	cys	Cylinder—short
1	1	0	1	1	cyl	Cylinder—long
1	1	0	1	0	t	Top
1	1	0	0	1	lp	Legs/prongs
1	1	0	0	0	slp	Single leg/prong
1	0	1	1	1	sph	Sphere
1	0	1	0	1	tap	Taper-to-point
1	0	1	0	0	cu	Curve
1	0	0	1	1	hsp	Half-sphere
1	0	0	0	1	pc	Plane—circular
1	0	0	0	0	r	Rim
0	1	1	1	1	br	Box—rectangular
0	1	1	1	0	pp	Parallel planes
0	1	1	0	1	bt	Box—thin
0	1	1	0	0	bl	Box—long
0	1	0	1	1	ps	Plane—square
0	1	0	1	0	pr	Plane—rectangle
0	1	0	0	1	lf	Long/flat
0	1	0	0	0	lft	Long/flat/thin
0	0	1	1	0	liq	Liquid
0	0	0	0	1	ho	Hole
0	0	0	0	0		(No second or third component)

[a]The meanings of the features are roughly (1) contains curves, (2) sides contain parallel lines, (3) sizes along all three dimensions of the same order of magnitude, (4) more equal in dimensions than shapes with similar values for preceding features, and (5) more regular than shapes with similar values for preceding features. The "code" letters will be used to describe the assignment of visual features to objects.

Acknowledgments

Most of the research presented in this paper was carried out while the authors were visiting scientists in the Departments of Psychology and Computer Science at the University of Toronto under the generous support and guidance of Geoff Hinton. We also wish to thank Marlene Behrmann, Martha Farah, Jay McClelland, Tom Mitchell, Dave Tourtezky, Rich Zemel, and an anonymous reviewer for their help. All of the simulations described in this paper were run on a Silicon Graphics Iris-4D/240S using an extended version of the Xerion simulator developed by Tony Plate. This research was supported by Grant 87–2–36 from the Alfred P. Sloan Foundation, Grant I89-01245-

Table 7. The Encoding of the Position and Size of the Second and Third Components Relative to the Main Component

Relative Position			
Features		*Code*	*Description*
1	1	ee	Extension at end
1	0	em	Extension at middle
0	1	ae	Attachment at end
0	0	am	Attachment at middle
0	0		(No second or third component)

Relative Size				
Features			*Code*	*Description*
1	1	1	l	Longer
1	1	0	e	Equal
0	1	1	s	Smaller
0	0	1	ms	Much smaller (1/4 to 1/2)
0	0	0	vms	Very much smaller
0	0	0		(No second or third component)

016 from the McDonnell-Pew Program in Cognitive Neuroscience, and Grant ASC-9109215 from the National Science Foundation

Reprint requests should be sent to Dr David C Plaut, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213-3890

Notes

1 Only one left hemisphere agnosic patient is known to us who makes predominantly visual errors, patient RV of Beauvois and Saillant (1985) As he was 75 when first tested, an additional apperceptive component is plausible

2 Our analysis does not, however, require us to assume a unitary semantics

3 In simulations reported by Plaut (1991), a network with a somewhat different architecture produced equivalent results as those described in this paper

4 See Figure 6 of Hinton and Shallice (1991) for a related empirical phenomena: semantic errors tend to have a smaller gap for a given proximity than do visual errors

5 This proportion is much higher than the average correct performance for subsequently presented objects after clean-up lesions (59 2%) because the prime is presented with all of the short-term weights set to zero

6 The explicit error rate of the patient (49/59) is much higher than in the simulation after V→I(0 5) lesions (3 5%), suggesting that either the patient used a lower response criterion or that naming involved random guessing In either case this would inflate the rate of "other" errors

Table 8. The Coding for General Visual Characteristics, Color, Texture, and Absolute Size

General Characteristics[a]	
Code	*Description*
dv	Direction of main component—vertical
dh	Direction of main component—horizontal
ss	Screw/sawtooth
ifl	Internal flexibility between components
con	Concave
sh	Sharp
dis	Distortable
int	Interior visible
dr	Rectangle/handle apparent on surface

Color				
Features			*Code*	*Description*
1	1	1	va	Various
1	1	0	brn	Brown
1	0	1	si	Silver
1	0	0	gr	Gray
0	0	1	wh	White
0	0	0	tr	Transparent

Texture			
Features		*Code*	*Description*
1	0	sm	Smooth
0	1	ro	Rough
0	0	ei	Either

Absolute Size						
Features					*Code*	*Description*
1	0	0	0	0	s<3i	Less than 3 in
1	1	0	0	0	s3–6i	3 to 6 in
1	1	1	0	0	s3–12i	3 to 12 in
0	1	1	0	0	s6–12i	6 to 12 in
0	1	1	1	0	s6i–2f	6 in to 2 ft
0	0	1	1	0	s1–2f	1 to 2 ft
0	0	0	1	1	s2–6f	2 to 6 ft
0	0	0	0	1	s>6f	Greater than 6 ft

[a]Each "general" characteristic is represented by a separate feature

Table 9. A Description of the Visual Representation of Each Object in Terms of the Codes Used in Previous Tables to Describe Values of Each Type of Visual Information

	Assignment of Visual Features to Objects										
Object	*Main*		*Second*			*Third*		*General*	*Color*	*Texture*	*Size*
cup	cyh	pc	ae	e	cu	am	s	dv con	va	sm	s3–6i
spoon	lf	hsp	ee	ms				con	si	sm	s3–12i
pan	cyh	ps	ae	e	lf	ae	l	dv con	si	sm	s6i–2f
fork	lf	lp	ee	ms				sh	si	sm	s3–12i
knife	lf	lft	ee	e				sh	si	sm	s3–12i
bowl	hsp							dv con	va	sm	s6i–2f
can	cy							dv	va	sm	s3–6i
plate	pc	r	ee	ms				dh	wh	sm	s6–12i
dish	pc	r	ae	ms				dh con	va	sm	s6–12i
glass	cyh	pc	ae	e				dv con	tr	sm	s3–6i
pen	cyl	tap	ee	vms	t	ee	ms	ifl	va	sm	s3–6i
file	pr	pr	ae	e				ifl	va	ei	sl–2f
paper	pr							dis	wh	sm	s6–12i
book	pr	pr	ae	e				ifl int	va	ei	s6–12i
disk	ps	ho	am	vms					bl	ro	s3–6i
tape	cys	ho	am	s					brn	sm	s3–6i
stamp	ps							dis	va	sm	s<3i
board	pr								va	sm	s1–2f
glue	lft	tap	ee	vms					va	sm	s<3i
ink	cyh	t	ee	e	liq	am	s	con int	va	sm	s<3i
chair	ps	lp	ae	e	ps	ae	e	dh	brn	ei	s2–6f
table	pr	lp	ae	s				dh	brn	sm	s2–6f
bed	bt	lp	ae	vms	pr	ae	ms	dh	va	ro	s>6f
sofa	br	pr	am	e	pp	ae	s	dh	va	ro	s2–6f
stool	pc	lp	ae	l				dh	brn	sm	s1–2f
rug	pc							dh dis	va	ro	s>6f
radio	bt							dh	va	sm	s6–12i
tele	pr	br	am	e				dh	gr	sm	s1–2f
divan	bt							dh	va	ro	s2–6f
desk	pr	pp	ae	s				dh dr	brn	sm	s2–6f
saw	lft	cu						ss sh	si	sm	s1–2f
nail	cyl	tap	ee	ms	pc	ee	l	sh	gr	sm	s3–6i
plane	bl	cu	am	ms	lft	am	vms	sh	gr	sm	s6i–2f
ruler	lf								brn	sm	s1–2f
screw	tap	pc	ee	l				ss sh	gr	ro	s<3i
awl	cyl	tap	ee	e	sph	ee	ms	sh	gr	sm	s3–6i
axe	bt	tap	ee	e	cyl	am	l	sh	gr	sm	s1–2f
bolt	cy	cys	ee	l				ss	gr	ro	s<3i
nut	cys	ho	am	s				ss	gr	ro	s<3i
vise	pp	cyl	am	e				dv ifl int	va	sm	s1–2f

Figure 10. Assignment of visual features to objects

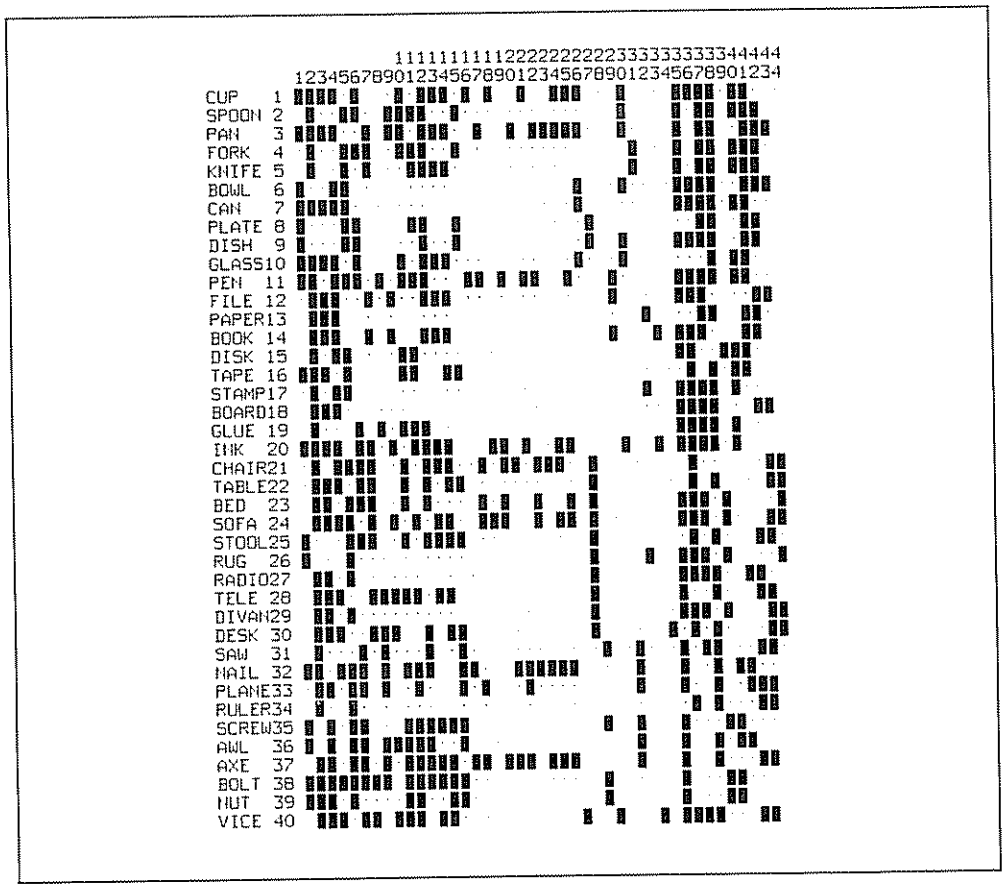

Figure 11. The assignment of semantic features to objects

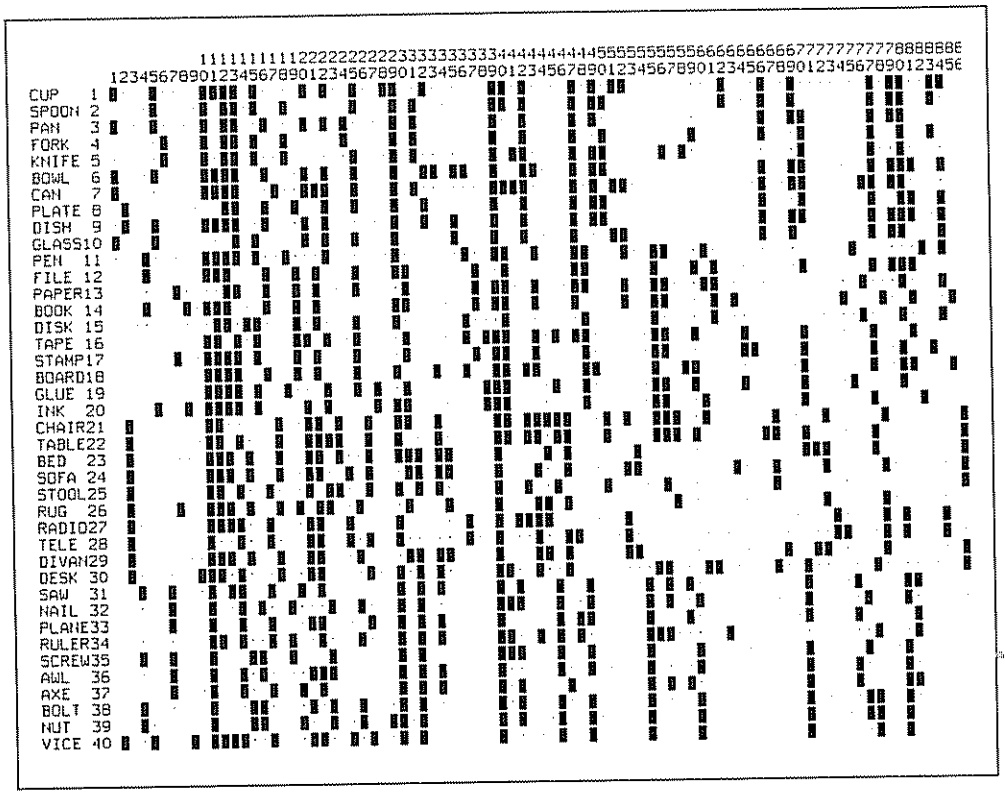

Table 10. The Semantic Features Used to Describe Objects

Semantic Features		
Visual Characteristics	*Made of*	*Specific Function*
1 Main component vertical	31 Metal	58 Chopping/cutting
2 Main component horizontal	32 Pottery	59 Holding in place
3 Screw/sawtooth	33 Wood	60 Writing
4 Internal flexibility	34 Cloth	61 Information-holding
5 Concave	35 Glass	62 Measuring
6 Sharp	36 Plastic	63 Reading
7 Distortable	37 Paper	64 Sticking
8 Interior visible	38 Other substance	65 Assigning-value
9 Rectangle/handle apparent	*Where Found*	66 Holding food/drink
10 Color (visual coding)	39 Home	67 Sitting
11 Color (visual coding)	40 Office	68 Lying
12 Color (visual coding)	41 Outdoors	69 Use-with-liquid
13 Smooth	42 Kitchen/dining room	70 Use-with-solid
14 Rough	43 Living room/study	71 Sleeping
15 Size less than 6 in	44 Bedroom	72 For comfort
16 Size 6 in to 2 ft	45 Work-room	73 For listening
17 Size greater than 2 ft	46 On ground	74 For viewing
18 Main-shape 1D	47 On surface	75 Manipulating another artifact
19 Main-shape 2D	48 Otherwise supported	76 Is manipulated by another artifact
20 Main-shape 3D	*General Function*	77 Functioning with another object
21 Rectangular cross section	49 Cooking	78 Functioning alone
22 Circular cross section	50 Eating	79 Container
23 Has legs	51 Drinking	*General Action*
24 Has other appendage	52 Leisure	80 Use with one arm
25 Simple	53 Rest	81 Use with two arms
26 Complex	54 Carpentry	82 Use with hand (little arm movement)
27 Liquid	55 Work-office	83 Use involves mouth
28 Has hole	56 Work-home	84 Easily breakable
Consistency	57 Aesthetic	85 Placed in lap/held in front of body
29 Hard		86 Characteristic action of whole body
30 Soft		

REFERENCES

Albert, M L, & Sandson, J (1986) Perseveration in aphasia *Cortex, 22,* 103–115

Allport, D A (1985) Distributed memory, modular systems and dysphasia In S K Newman & R Epstein (Eds.), *Current perspectives in dysphasia* Edinburgh: Churchill Livingstone

Assal, G, & Regli, F (1980) Syndrome de disconnection visuo-verbale et visuo-gesturelle *Revue Neurologique, 136,* 365–376

Atiya, A, & Baldi, P. (1989) Oscillations and synchronizations in neural networks: An exploration of the labeling hypothesis *International Journal of Neural Systems, 1*(2), 103–124.

Baldi, P, & Meir, R (1990) Computing with arrays of coupled oscillators: An application to preattentive texture discrimination. *Neural Computation, 2*(4), 459–471

Barnes, C A, & McNaughton, B L (1980). Spatial memory and hippocampal synaptic plasticity in middle-aged and senescent rats In D Stein (Ed), *The Psychobiology of Aging: Problems and Perspectives,* (pp 253–272) New York: Elsevier

Beauvois, M F (1982) Optic aphasia: A process of interaction between vision and language *Proceedings of the Royal Society of London, Series B,* 298, 35–47

Beauvois, M F., & Saillant, B (1985) Optic aphasia for colours and colour agnosia: A distinction between visual and visuo-verbal impairments in the processing of colours *Cognitive Neuropsychology, 2*(1), 1–48

Biederman, I (1987) Recognition-by-components: A theory of human image understanding *Psychological Review, 94,* 115–147

Bliss, T V P, & Lømo, T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of unanaesthetized rabbit following stimulation of the perforant path *Journal of Physiology, 232,* 331–356

Bobick, A (1987) *Natural object categorization* Ph D thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology

Bush, P C, & Douglas, R J (1991) Synchronization of bursting action potential discharge in a model network of neocortical neurons. *Neural Computation, 3,* 19–30

Caramazza, A, Hillis, A., Rapp, B C., & Romani, C (1990) The multiple semantics hypothesis: Multiple confusions? *Cognitive Neuropsychology, 7,* 161–189

Churchland, P S, & Sejnowski, T J (1988) Perspectives on cognitive neuroscience. *Science, 242,* 741–745

Cleeremans, A, & McClelland, J L (1991) Learning the structure of event sequences *Journal of Experimental Psychology: General, 120,* 235–253.

Collins, A M, & Quillian, M R (1969). Retrieval time from semantic memory *Journal of Verbal Learning and Verbal Behaviour, 8,* 240–248

Coltheart, M (1980) Deep dyslexia: A right-hemisphere hypothesis In M Coltheart, K E Patterson, & J C Marshall (Eds), *Deep dyslexia* (pp 326–380) London: Routledge & Kegan Paul

Coltheart, M (1983) The right hemisphere and disorders of reading In A Young (Ed), *Functions of the right cerebral hemisphere* New York: Academic Press

Coltheart, M, Patterson, K E, & Marshall, J C (1987) Deep dyslexia since 1980 In M. Coltheart, K E. Patterson, & J C Marshall (Eds), *Deep dyslexia* (pp 407–451) London: Routledge & Kegan Paul

Coslett, H B, & Saffran, E M. (1989) Preserved object recognition and reading comprehension in optic aphasia *Brain, 112,* 1091–1110

Cotman, C W, & Monaghan, D T (1988) Excitatory amino acid neurotransmission: NMDA receptors and Hebb-type synaptic plasticity *Annual Review of Neuroscience, 11,* 61–80

Coughlan, A K, & Warrington, E K (1978). Word-comprehension and word-retrieval in patients with localised cerebral lesions. *Brain, 101,* 163–185

Crick, F (1984) The function of the thalamic reticular complex: The searchlight hypothesis *Proceedings of the National Academy of Science, U S A, 81,* 4586

Crick, F (1989) The recent excitement about neural networks *Nature (London), 337,* 129–132

Crick, F H C, & Asanuma, C. (1986) Certain aspects of the anatomy and physiology of the cerebral cortex In J L McClelland, D E. Rumelhart, & the PDP research group (Eds), *Parallel distributed processing: Explorations in the microstructure of cognition Volume 2: Psychological and biological models* (pp 333–371) Cambridge, MA: MIT Press

Dehaene, S, & Changeux, P (1991) The Wisconsin Card Sorting test: Theoretical analyses and modelling in a neuronal network *Cerebral Cortex, 1,* 62–79

Eckhorn, R, Bauer, R, Jordan, W, Brosh, M, Kruse, W, Munk, M, & Reitboeck, H J (1988) Coherent oscillations: A mechanism of feature linking in the visual cortex? Multiple electrode and correlation analysis in the cat *Biological Cybernetics, 60,* 121–130

Eckhorn, R, Reitboeck, H J, Ardnt, M, & Dicke, P (1989) A neural network for feature linking via synchronous activity: Results from cat visual cortex and from simulations In R M J Cottrill (Ed), *Models of brain function* Cambridge, England: Cambridge University Press

Engel, A K, König, P, Gray, C M, & Singer, W (1990) Stimulus-dependent neuronal oscillations in cat visual cortex: Inter-columnar interaction as determined by cross-correlation analysis. *European Journal of Neuroscience, 2,* 586–606

Farah, M J (1990) *Visual Agnosia: Disorders of object recognition and what they tell us about normal vision* Cambridge, MA: MIT Press

Farah, M J (1991) Patterns of co-occurrence among the associative agnosias: Implications for visual object representation *Cognitive Neuropsychology, 8*(1), 1–19

Farah, M J, & McClelland, J L (1991) A computational model of semantic memory impairment: Modality-specificity and emergent category-specificity *Journal of Experimental Psychology: General, 120*(4), 339–357

Feldman, J A (1982) Dynamic connections in neural networks *Biological Cybernetics, 46,* 27–39.

Frith, C D, Friston, K, Liddle, P F, & Frackowiak, R S J (1991) Willed action and the prefrontal cortex in man *Proceedings of the Royal Society of London, Series B, 244,* 241–246

Gardner-Medwin, A R (1989) Doubly modifiable synapses: A model of short and long term auto-associative memory. *Proceedings of the Royal Society of London, Series B, 238,* 137–154

Geshwind, N (1965) Disconnexion syndromes in animals and man *Brain, 88,* 237–294

Gibson, J J (1979) *The ecological approach to visual perception* Boston: Houghton Mifflin

Gil, R, Pluchon, C, Toullat, G, Michenau, D, Rogez, R, & Levevre, J P (1985). Disconnexion visuo-verbale (aphasie optique) pour les objets, les images, les couleurs et les visages avec alexie "abstractive." *Neuropsychologia, 23,* 333–349

Goebel, R (1990) Binding, episodic short-term memory, and selective attention, or Why are PDP models so poor at symbol manipulation? In D S Touretzky, J L Elman, T J Sejnowski, & G E Hinton (Eds), *Proceedings of the 1990*

Connectionist Models Summer School San Mateo, CA: Morgan Kaufmann.
Goodglass, H., Barton, M., & Kaplan, E. (1968). Sensory modality and object naming in aphasia. *Journal of Speech and Hearing Research, 3,* 257–267.
Gordon, B. (1982). Confrontation naming: Computational model and disconnection simulation. In M. A. Arbib, D. Caplan, & J. C. Marshall (Eds.), *Neural models of language processes.* New York: Academic Press.
Gray, C. M., Konig, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflect global stimulus properties. *Nature (London), 338,* 334–337.
Hartzell, H. C. (1981). Mechanisms of slow synaptic potentials. *Nature (London), 291,* 539–543.
Hécaen, H., & Angelergues, R. (1963). *La cécité psychique.* Paris: Masson & Cie.
Heilman, K. M., & Valenstein, E. (1985). *Clinical neuropsychology.* Oxford: Oxford University Press.
Helmick, J., & Berg, C. (1976). Perseveration in brain-injured adults. *Journal of Communication Disorders, 9,* 571–582.
Hinton, G. E. (1989). Connectionist learning procedures. *Artificial Intelligence, 40,* 185–234.
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: Lawrence Erlbaum.
Hinton, G. E., & Shallice, T. (1989). *Lesioning a connectionist network: Investigations of acquired dyslexia.* Technical Report CRG-TR-89-3, Department of Computer Science, University of Toronto, Toronto, Ontario.
Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review, 98*(1), 74–95.
Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain.*
Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Science, U.S.A., 79,* 2554–2558.
Horn, D., Sagi, D., & Usher, M. (1991). *Segmentation, binding and illusory conjunctions.* Technical Report CS91-07, Department of Applied Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel.
Howard, D., & Franklin, S. (1988). *Missing the meaning?* Cambridge, MA: MIT Press.
Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review.*
Humphreys, G. W., & Riddoch, M. J. (1987). *To see but not to see: A case-study of visual agnosia.* Hillsdale, NJ: Lawrence Erlbaum.
Iorio, L., Falanga, A., Fragassi, N. A., & Grossi, D. (1992). Visual associative agnosia and optic aphasia: A single case study and review of the syndromes. *Cortex, 28,* 23–37.
Joordens, S., & Besner, D. (1992). "Priming" effects that span an intervening unrelated word: Implications for models of memory representation and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(3), 483–491.
Kammen, D., Koch, C., & Holmes, P. J. (1990). Collective oscillations in the visual cortex. In D. S. Tourtezky (Ed.), *Advances in Neural Information Processing Systems 2* (pp. 76–83). San Mateo, CA: Morgan Kaufmann.
Konig, P., & Shillen, T. B. (1991). Stimulus-dependent assembly formation of oscillatory responses: I. Synchronization. *Neural Computation, 3*(2), 155–166.
Kupferman, I. (1979). Modulatory actions of neurotransmitters. *Annual Review of Neuroscience, 2,* 447–465.
Larrabee, G. J., Levin, H. S., Huff, F. J., Kay, M. C., & Guinto, F. C. (1985). Visual agnosia contrasted with visual verbal disconnection. *Neuropsychologia, 23*(1), 1–12.
Levine, D. N. (1978). Prosopagnosia and visual object agnosia: A behavioral study. *Brain and Language, 5,* 341–365.
Levine, D. S., & Prueitt, P. S. (1989). Modeling some effects of frontal lobe damage—Novelty and perseveration. *Neural Networks, 2,* 103–116.
Lhermitte, F., & Beauvois, M. F. (1973). A visual-speech disconnexion syndrome: Report of a case with optic aphasia, agnosic alexia and colour agnosia. *Brain, 96,* 695–714.
Lynch, G., McGaugh, J. L., & Weinberger, N. M. (1984). *Neurobiology of Learning and Memory.* New York: Guilford Press.
Lytton, W. W., & Sejnowski, T. J. (1991). Simulations of cortical pyramidal neurons synchronized by inhibitory interneurons. *Journal of Neurophysiology, 66,* 1–22.
Mandler, G. (1980). Recognizing: The judgement of previous occurrence. *Psychological Review, 87,* 252–271.
Margolin, D. I., Marcel, A. J., & Carlson, N. R. (1985). Common mechanisms in dysnomia and post-semantic surface dyslexia: Processing deficits and selective attention. In K. E. Patterson, M. Coltheart, & J. C. Marshall (Eds.), *Surface dyslexia* (pp. 139–174). Hillsdale, NJ: Lawrence Erlbaum.
Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, Series B, 200,* 269–294.
McCarthy, R., & Warrington, E. K. (1986). Visual associative agnosia: A clinico-anatomical study of a single case. *Journal of Neurology, Neurosurgery and Psychiatry, 49,* 1233–1240.
McCarthy, R., & Warrington, E. K. (1990). *Cognitive neuropsychology.* New York: Academic Press.
McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86,* 287–330.
McClelland, J. L., & Kawamoto, A. H. (1986). Mechanisms of sentence processing: Assigning roles to constituents of sentences. In J. L. McClelland, D. E. Rumelhart, & the PDP research group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models* (pp. 272–325). Cambridge, MA: MIT Press.
McClelland, J. L. & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*(5), 375–407.
McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114*(2), 159–188.
McLaren, I. P. L., Kaye, H., & Mackintosh, N. J. (1989). An associative theory of the representation of stimuli: Applications to perceptual learning and latent inhibition. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neurobiology.* Oxford: Oxford University Press.
McNaughton, B. L. (1982). Long-term synaptic enhancement and short-term post-tetanic potentiation in rat fascia dentata act through different mechanisms. *Journal of Physiology, 324,* 249–262.
McNaughton, B. L. (1983). Activity dependent modulation of hippocampal synaptic efficacy: Some implications for memory processes. In W. Siefert (Ed.), *Neurobiology of the hippocampus* (pp. 233–252). New York: Academic Press.
McNaughton, B. L. & Morris, R. G. M. (1987). Hippocampal

synaptic enhancement and information storage within a distributed memory system *Trends in Neurosciences, 10*(10), 408–415

Meyer, D E, & Schvaneveldt, R W (1976) Meaning, memory structure, and mental processes *Science, 192*, 27–33

Milner, B (1963) Effects of different brain lesions on card-sorting *Annual Review of Neuroscience, 9*, 90–100

Milner, B (1964) Some effects of frontal lobectomy in man In J M Warren, & K Akert, K (Eds), *The frontal granular cortex and behavior* New York: McGraw-Hill

Morton, J (1969) The interaction of information in word recognition *Psychological Review, 76*, 165–178

Morton, J, & Patterson, K (1980) A new attempt at an interpretation, or, an attempt at a new interpretation. In M Coltheart, K E Patterson, & J C Marshall, J C (Eds), *Deep dyslexia* (pp 91–118) London: Routledge & Kegan Paul.

Nadel, L, Cooper, L A, Culicover, P, & Harnish, R M. (1989) *Neural connections, mental computation* Cambridge, MA: MIT Press

Newcombe, F, & Marshall, J C (1980) Transcoding and lexical stabilization in deep dyslexia In M Coltheart, K E Patterson, & J C. Marshall (Eds), *Deep dyslexia* (pp 176–188) London: Routledge & Kegan Paul

Palmer, S E (1977) Hierarchical structure in perceptual representation *Cognitive Psychology, 9*, 441–474

Patterson, K E, & Besner, D (1984) Is the right hemisphere literate? *Cognitive Neuropsychology, 3*, 341–367

Peña-Casanova, J, & Roig-Rovira, T (1985) Optic aphasia, optic apraxia, and loss of dreaming *Brain and Language, 26*, 63–71

Petersen, S E, Fox, P T, Posner, M I, Mintun, M, & Raichle, M E (1988) Positron emission tomographic studies of the cortical anatomy of single-word processing *Nature (London), 331*, 585–589

Plaut, D C (1991). *Connectionist neuropsychology The breakdown and recovery of behavior in lesioned attractor networks* Ph D thesis, School of Computer Science, Carnegie Mellon University Available as Technical Report CMU-CS-91-185

Plaut, D C, & Farah, M J (1990) Visual object representation: Interpreting neurophysiological data within a computational framework *Journal of Cognitive Neuroscience, 2*(4), 320–343

Plaut, D C, & Shallice, T (1993) Deep dyslexia: A case study of connectionist neuropsychology *Cognitive Neuropsychology*, in press.

Plaut, D C, & Shallice, T (1991a) *Deep dyslexia: A case study of connectionist neuropsychology* Technical Report CRG-TR-91-3, Connectionist Research Group, Department of Computer Science, University of Toronto, Toronto, Ontario

Plaut, D C, & Shallice, T. (1991b) Effects of abstractness in a connectionist model of deep dyslexia. In *Proceedings of the 13th Annual Conference of the Cognitive Science Society* (pp 73–78) Hillsdale, NJ: Lawrence Erlbaum

Poeck, K, & Luzzatti, C (1988) Slowly progressive aphasia in three patients: The problem of accompanying neuropsychological deficit *Brain, 111*, 151–168.

Ratcliff, G., & Newcombe, F A (1982) Object recognition: Some deductions from the clinical evidence In A W Ellis (Eds), *Normality and pathology in cognitive functions* (pp 147–171) New York: Academic Press

Richards, W (1988) *Natural computation* Cambridge, MA: MIT Press

Riddoch, M J, & Humphreys, G W (1987) Visual object processing in optic aphasia: A case of semantic access agnosia *Cognitive Neuropsychology, 4*(2), 131–185

Riddoch, M J, Humphreys, G W, Coltheart, M, & Funnell, E (1988) Semantic systems or system? Neuropsychological evidence re-examined *Cognitive Neuropsychology, 5*(1), 3–25

Rubens, A R (1979) Agnosia In K M Heilman, & E Valenstein (Eds), *Clinical neuropsychology* (pp 233–267) Oxford: Oxford University Press

Rumelhart, D E, Hinton, G E, & Williams, R J (1986) Learning representations by back-propagating errors *Nature (London), 323*(9), 533–536

Sacchett, C., & Humphreys, G W (1992) Calling a squirrel a squirrel but a canoe a wigwam: A category-specific deficit for artifactual objects and body parts *Cognitive Neuropsychology, 9*(1), 73–86

Saffran, E M, Bogyo, L C, Schwartz, M F, & Marin, O S M (1980) Does deep dyslexia reflect right-hemisphere reading? In M Coltheart, K E Patterson, & J C Marshall (Eds), *Deep dyslexia* (pp 381–406) London: Routledge & Kegan Paul

Sandson, J, & Albert, M L (1984) Varieties of perseveration *Neuropsychologia, 22*, 715–732

Schmidhuber, J (1992) Learning to control fast-weight memories: An alternative to dynamic recurrent networks *Neural Computation, 4*, 131–139

Schwartz, M F, Saffran, E M, & Marin, O S M (1980) Fractioning the reading process in dementia: Evidence for word-specific print-to-sound associations In M Coltheart, K E Patterson, & J C Marshall (Eds), *Deep dyslexia* (pp 259–269) London: Routledge & Kegan Paul

Shallice, T (1982) Specific impairments of planning *Proceedings of the Royal Society of London, Series B, 298*, 199–209

Shallice, T (1987) Impairments of semantic processing: Multiple dissociations. In M Coltheart, G Sartori, & R Job (Eds), *The cognitive neuropsychology of language* (pp 111–128) Hillsdale, NJ: Lawrence Erlbaum

Shallice, T (1988a) *From neuropsychology to mental structure* Cambridge: Cambridge University Press

Shallice, T (1988b) Specialisation within the semantic system *Cognitive Neuropsychology, 5*, 133–142

Shallice, T (1993) Multiple semantics: Whose confusions? *Cognitive Neuropsychology*, in press

Shallice, T., & Coughlan, A K (1980) Modality specific word comprehension deficits in deep dyslexia *Journal of Neurology, Neurosurgery and Psychiatry, 43*, 866–872

Shimamura, A P (1986) Priming effects in amnesia: Evidence for a dissociable memory function *Quarterly Journal of Experimental Psychology, 38A*, 619–644

Small, S L (1991) Focal and diffuse lesions of cognitive models In *Proceedings of the 13th Annual Conference of the Cognitive Science Society* (pp 85–90) Hillsdale, NJ: Lawrence Erlbaum

Smolensky, P. (1986) Neural and conceptual interpretation in PDP models. In J L McClelland, D E Rumelhart, & the PDP research group (Eds), *Parallel distributed processing Explorations in the microstructure of cognition. Volume 2 Psychological and biological models* (pp 390–431) Cambridge, MA: MIT Press, Cambridge

Smolensky, P (1988) On the proper treatment of connectionism *Behavioral and Brain Sciences, 11*, 1–74

Sompolinsky, H., Golomb, D, & Kleinfeld, D (1989). Global processing of visual stimuli in a neural network of coupled oscillators *Proceedings of the National Academy of Science, U S A, 87*, 7200–7204

Sporns, O, Gally, J A, Reeke, G N, & Edelman, G M (1989) Reentrant signaling among simulated neuronal groups leads to coherency in their oscillatory activity *Proceedings of the National Academy of Science, U S A, 86*, 7265–7269

Tyrrell, P J, Warrington, E K, Frackowiak, R S J, & Rossor,

M N (1990) Heterogeneity in progressive aphasia due to focal cortical atrophy: A clinical and PET study *Brain, 113,* 1321–1336.

von der Malsburg, C (1981) *The correlation theory of brain function* Technical Report 81–2, Department of Neurobiology, Max-Plank-Institute for Biophysical Chemistry, Gottingen, Germany

von der Malsburg, C. (1988). Pattern recognition by labeled graph matching *Neural Networks, 1,* 141–148

von der Malsburg, C, & Schneider, W (1986) A neural cocktail-party processor. *Biological Cybernetics, 54,* 29–40.

Warrington, E K, & Shallice, T (1984) Category specific semantic impairments *Brain, 107,* 829–853

Warrington, E K, & Taylor, A M (1973) The contribution of the right parietal lobe to object recognition *Cortex, 9,* 152–164

Warrington, E K, & Taylor, A M (1978) Two categorical stages of object recognition. *Perception, 7,* 695–705.

Williams, R J., & Peng, J (1990) An efficient gradient-based algorithm for on-line training of recurrent network trajectories. *Neural Computation, 2*(4), 490–501

Wilson, M W., & Bower, J M (1990) Computer simulation of oscillatory behavior in cerebral cortical networks In D S Tourtezky (Ed), *Advances in Neural Information Processing Systems 2* (pp 84–91) San Mateo, CA: Morgan Kaufmann