COGNITIVE NEUROPSYCHOLOGY, 2008, 25 (2), 136–164

Psychology Press
Taylor & Francis Group

A single-system account of semantic and lexical deficits in five semantic dementia patients

Katia Dilkina
Carnegie Mellon University, Pittsburgh, PA, USA, and Stanford University, Stanford, CA, USA

James L. McClelland
Stanford University, Stanford, CA, USA

David C. Plaut
Carnegie Mellon University, Pittsburgh, PA, USA

In semantic dementia (SD), there is a correlation between performance on semantic tasks such as picture naming and lexical tasks such as reading aloud. However, there have been a few case reports of patients with spared reading despite profound semantic impairment. These reports have sparked an ongoing debate about how the brain processes conceptual versus lexical knowledge. One possibility is that there are two functionally distinct systems in the brain—one for semantic and one for lexical processing. Alternatively, there may be a single system involved in both. We present a computational investigation of the role of individual differences in explaining the relationship between naming and reading performance in five SD patients, among whom there are cases of both association and dissociation of deficits. We used a connectionist model where information from different modalities feeds into a single integrative layer. Our simulations successfully produced the overall relationship between reading and naming seen in SD and provided multiple fits for both association and dissociation data, suggesting that a single, cross-modal, integrative system is sufficient for both semantic and lexical tasks and that individual differences among patients are essential in accounting for variability in performance.

Keywords: Semantic processing; Lexical processing; Semantic dementia; Connectionist modelling.

Is lexical knowledge separate from other kinds of knowledge (sensory-motor, encyclopaedic, etc.)? Is it represented or processed differently in the brain? Ultimately, these kinds of questions come down to the topic of a "mental lexicon"—a theoretical construct that includes detailed orthographic, phonological, and morpho-syntactic knowledge about words. Essentially, a mental lexicon stores and organizes a person's vocabulary. While lexical knowledge can be viewed as knowledge of words, semantic or conceptual knowledge is the knowledge of things, which are in turn

Correspondence should be addressed to Katia Dilkina, Jordan Hall (Bldg 420), Stanford, CA, 94305–2130, USA (E-mail: knd@andrew.cmu.edu).

Supported by National Institute of Mental Health (NIMH) Interdisciplinary Behavioral Science Center Grant MH 64445 (JLM, Program Director). We thank Tim Rogers and Matt Lambon Ralph for their helpful discussions and comments on earlier versions of this paper.

http://www.psypress.com/cogneuropsychology DOI:10.1080/02643290701723948

named or described with words. For example, one's knowledge of the word "dog", how to spell and pronounce it, and how to make its plural, is lexical knowledge, while knowing what dogs look and sound like, how they behave, and how to interact with them is semantic knowledge. The core of the issue is whether lexical knowledge and semantic knowledge are represented separately in the brain.

A relevant set of data that speaks to this question comes from semantic dementia patients. Semantic dementia (SD) is a selective impairment of conceptual knowledge (Hodges, Patterson, Oxbury, & Funnell, 1992; Snowden, Goulding, & Neary, 1989) due to progressive atrophy of the anterior inferior and lateral aspects of the temporal cortex (Galton et al., 2001; Mummery et al., 2000). The atrophy typically starts from the pole and progresses posteriorly. It is usually bilateral but often asymmetric. Tasks that present difficulty to SD patients span input and response modalities: They include object and picture naming (e.g., Hodges, Graham, & Patterson, 1995), smell naming (Luzzi et al., 2007), object recognition from touch (Coccia, Bartolini, Luzzi, Provinciali, & Lambon Ralph, 2004), smell-to-picture matching (Luzzi et al., 2007), word-to-picture matching, category matching, delayed copying (e.g., Lambon Ralph & Howard, 2000), matching environmental sounds to pictures (Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000), and nonverbal tests of semantic association such as the picture versions of the Pyramids and Palm Trees (PPT) test (Howard & Patterson, 1992) and the Camel and Cactus test (Bozeat et al., 2000). Impaired performance on these and similar tasks and findings of association and item consistency across tasks (Bozeat et al., 2000; Luzzi et al., 2007) indicate a cross-modal semantic deficit (see also Garrard & Carroll, 2006).

Not only do SD patients consistently score very poorly, but the majority of overt errors they make are classifiable as "semantic" (Hodges et al., 1995; Rogers et al., 2004a). For example, in naming, calling a tiger "a lion" is a semantic error (which can be contrasted with a phonological error, for example, "a timer", or an unrelated error, e.g., "a piano"). Semantic errors often involve choosing an item from the same category as the target or an item closely associated with the target (e.g., bed instead of hammock), even in the absence of visual similarity between the target and the foil (e.g., pliers vs. hammer). Remarkably, despite the semantic deficit, other cognitive abilities such as memory for recent episodes, nonverbal reasoning, visuo-spatial abilities, working memory, phonology, and syntax are largely spared (Patterson & Hodges, 2000; Snowden, Neary, & Mann, 1996).

Notably, SD patients are also impaired on a number of tasks that are traditionally considered lexical in that they tap knowledge about words other than the word's meaning. These include word reading (Funnell, 1996; Patterson & Hodges, 1992), word spelling (Graham, Patterson, & Hodges, 2000), verb past-tense inflection (Cortese, Balota, Sergent-Marshall, Buckner, & Gold, 2006; Patterson, Lambon Ralph, Hodges, & McClelland, 2001), and lexical decision (Rogers, Lambon Ralph, Hodges, & Patterson, 2004b). The deficit is most prominent for atypical low-frequency items. The patients show surface dyslexia and surface dysgraphia—inability to read and spell irregular (especially low-frequency) words. Most errors are legitimate alternative rendering of components (LARC) errors (Patterson, Suzuki, Wydell, & Sasanuma, 1995), a large subset of which are regularizations (Funnell, 1996; Patterson & Hodges, 1992; Woollams, Lambon Ralph, Plaut, & Patterson, 2007). A similar pattern is also seen in verb past-tense inflection (Patterson et al., 2001).

In the vast majority of patients, there is an association between semantic deficits and impairment on lexical tasks such as word reading. A recent paper by Patterson et al. (2006) reported an investigation of the relationship between semantic proficiency and each of four lexical tasks and two nonverbal tasks. A total of 14 semantic dementia patients were tested using the same testing battery. A strong positive correlation was found between the patients' composite semantic score and their performance on atypical items in each of the tasks. Such findings have motivated the idea that impairment on all of these tasks arises as

a consequence of damage to a single integrated system that mediates both semantic and lexical processing.

This association seen in the patient data has been accounted for in a number of connectionist models (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996; Rogers et al., 2004a; Seidenberg & McClelland, 1989) with an overall architecture as depicted in Figure 1. Within this theory, damage to the integrative layer called "semantics" leads to disruption of naming and verbal definition, since the integrative layer mediates between the visual input and the phonological output. Damage to the integrative layer also leads to difficulties in reading. This is true even though the model provides a "direct route" from orthography and phonology. As discussed in Plaut et al. (1996), when reading is learned in a system with the architecture shown in Figure 1, a division of labour develops, such that the direct pathway becomes particularly effective at pronouncing items that are high in either frequency or spelling–sound consistency, including pronounceable nonwords, since it is sensitive to the systematicity in the mapping between spelling and sound. The pathway mediated by the integrative layer, on the other hand, must learn the largely arbitrary mappings between spelling and various types of semantic information (including, for example, what a DOG looks like) and so is less sensitive to this systematicity. It therefore comes to play an especially important role in reading exceptional items, particularly those of low frequency. Although all words draw on both pathways, and items of high frequency tend to be robustly encoded in both pathways, regular items of low frequency can still be processed by the direct pathway. Exceptions will often be regularized by this pathway acting alone, so that the input from the integrative layer is crucial for reading such items. As the integrative layer is damaged more and more, the system becomes impaired on reading irregular (especially low-frequency) items, because they are the items that most strongly rely on semantic support. Thus, this framework is able to account for the observed association between impairment on tasks such as naming and PPT and deficits such as surface dyslexia and dysgraphia.

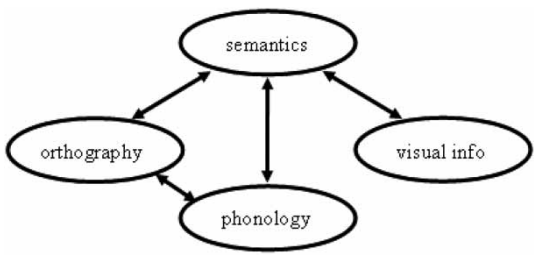

Figure 1. *Generic parallel distributed processing (PDP) model of semantic and lexical processing.*

It must be noted, however, that in addition to the growing collection of reports of patients with associations between semantic and lexical processing, there have been a few case studies of individual patients showing dissociation—patients W.L.P. (Schwartz, Saffran, & Marin, 1980), E.M. (Blazely, Coltheart, & Casey, 2005), and D.R.N. (Cipolotti & Warrington, 1995) showed little or no impairment on reading despite their profound semantic deficits. E.M. showed impairment on picture naming (34% correct), name comprehension (65% correct), and word–picture matching (written: 63% correct; spoken: 62% correct), and very little impairment on word reading (98% correct) and lexical decision (97% correct). Similarly, D.R.N. performed remarkably well on reading (98% overall performance averaged over four independent tests). In addition to reading, she was only tested on naming and verbal definition, and showed severe impairment on both tasks (25% correct on naming and 52% overall performance on verbal definition averaged over four tests). Finally, when patient W.L.P. was first tested, despite her very poor performance on word–picture matching (15% correct) and category sorting (60% correct averaged over three categories), she did exceptionally well (98% correct) on reading a list of high-frequency words, half of which had atypical spelling-to-sound correspondence. Case reports of SD patients without surface dyslexia have been rare. We are aware of only 6 such patients, while there have been an overwhelming number of individuals who do

show a reading disability, 48 such cases reported by Woollams et al. (2007) alone. Furthermore, when such individuals have been followed longitudinally, a surface dyslexia pattern eventually ensues. For example, in the case of W.L.P., only six months after the testing session described above, her reading performance on the same set of items dropped to 85% and, four months after that, to 77%.

A recent study documenting the reading performance of seven SD patients found that even when reading may appear intact in terms of accuracy, the reading latencies of the patients are much larger than those of age-, education-, and occupation- matched controls, suggesting that whatever mechanism may underlie reading, it has been compromised in these individuals (McKay, Castles, Davis, & Savage, 2007). Just like the rest of the patients in this group, patient R.R., whose reading accuracy was within the normal range, also showed significantly slower reaction times than controls. In addition, the longer latencies were associated with items that showed impairment in the semantic tasks, supporting the notion that the reading mechanism is not divorced from the semantic system (McKay et al., 2007).

Despite these findings and the rarity of SD patients with spared reading, such cases have been used as support for the argument that the semantic and the lexical deficits—though often co-occurring—are in fact distinct deficits and are caused by neurological damage to two functionally distinct (even if anatomically neighbouring) systems (e.g., Coltheart, 2004).

The present study includes a set of computational simulations investigating the possibility that it is unnecessary to postulate separate lexical and semantic systems. We adopt a single-system perspective in which information of different types (e.g., what an object looks like, what it is called, how one interacts with it) and from different modalities (visual, auditory, etc.) is integrated. We suggest that the cases of association and the cases of dissociation may lie on a continuum of performance that is shaped by individual differences existing prior to brain damage—including both biological differences based on genetic factors and differences in experience—and individual differences in the severity and spatial distribution of the progressive brain damage within the semantic system. In this approach, the different SD patients are seen as coming from a single distribution and falling at different points within that distribution rather than as individual cases somehow fundamentally different from each other. A similar approach has been taken by Dell and colleagues (Dell, Martin, & Schwartz, 2007; Schwartz, Dell, Martin, Gahl, & Sobel, 2006) to account for the picture-naming profiles of a large group of aphasic patients. Variability in the patient performance and error patterns was modelled in their interactive two-step model of lexical access by lesioning semantic and/or phonological connections. The model was then used to successfully predict how the same set of patients perform on word repetition.

The aim of the current project was to shed more light on the reasons why performance on naming and reading may be partially but not perfectly correlated across SD patients. Within a single integrative system, the robustness of performance on the different tasks may depend on a number of factors, so that the observed differences in performance in SD patients might arise from any (or a combination of) individual differences.

The idea that premorbid individual differences may play an important role in accounting for SD patients' performance on lexical tasks has been already suggested by Plaut (1997), where a parallel distributed processing (PDP) model was used to show how the competence of the direct pathway (i.e., the pathway between orthography and phonology, which does not rely on semantics and is therefore crucial for nonword reading) could vary drastically depending on its learning properties and the strength of the semantic contribution to phonology. Differences in the division of labour between the direct and the semantic pathways of reading could account for the variability in the data—since semantic damage does not affect the direct pathway.

The importance of individual differences in explaining the reading performance of semantic dementia patients has recently been supported by

an extensive report documenting 100 observations of reading data from 51 patients (Woollams et al., 2007). Similarly to Patterson et al. (2006), the authors found a strong correlation between the patients' semantic impairment and their reading deficit. The very occasional cases showing semantic deterioration but relatively preserved reading ability as well as the opposite pattern of unusually impaired reading could be accounted for by positing individual differences in premorbid reliance on semantics during reading. This was demonstrated in simulations of a connectionist network using the architecture and stimuli of Plaut et al. (1996), following the methods used in Plaut (1997). Also, longitudinally, all cases with initially spared reading inevitably showed the expected reading deficit as the semantic impairment worsened. Their longitudinal profiles paralleled the longitudinal profile of the group mean, showing that individual differences were relatively stable over time.

In the current study, we further explore the individual differences hypothesis and extend this work to address several important issues. Firstly, the Plaut et al. model provided only an in-principle argument about the role of semantics. It did not implement semantics, but instead, provided an input to the network's phonological layer with characteristics presumed to mimic those that would actually be generated by semantics. Here we implemented semantics in the form of an integrative layer of hidden units mediating orthographic, phonological, visual, and motor/action information. Second, the Plaut et al. model only simulated reading while the architecture of the present model allows us to simulate both reading and naming. Finally, the Plaut et al. (1996) and Plaut (1997) simulations involved manipulating weight decay and semantic strength in the network; it is not clear how these parameters are grounded in actual characteristics of the population. In contrast, the model we describe here allows us to explore network parameters that are explicitly related to potentially measurable individual difference variables.

Specifically, the parallel distributed processing model presented here implemented and manipulated three individual differences: (a) differences in reading experience (i.e., training regime in the network); (b) differences in the neural pathway mapping orthography to phonology (i.e., direct pathway size in the network); and/or (c) differences in the spatial distribution of the atrophy (i.e., lesion distribution bias in the network). We now discuss the motivation for each of these three manipulations.

Motivation for the reading experience manipulation

Nationwide annual surveys show that there is large variability in all age groups in the amount that people read and in their literacy skills. According to a 2002 survey of the reading habits of Americans above 25 years old, only 47.3% of the population reported reading any literary piece in the past year (Rooney et al., 2006, p. 143). The results from the 2003 National Assessment of Adult Literacy (NAAL) survey indicated that about a third (31.6%) of Americans above 16 years old read books every day, and another third (37.9%) do so less than once a week, or never (Rooney et al., 2006, p. 154).

Given this wide range of reading experience, it is interesting to consider the two SD patients reported by Blazely et al. (2005). One of these patients, E.M., was a secretary who had completed high school and a secretarial course, while the other, P.C., was an air conditioning salesman who had not even completed high school. Based on their education and occupations, and in the context of the findings reported above, it seems plausible that these two individuals had significantly different amounts of reading and/or spelling experience in their lives, which probably resulted in different premorbid competence levels of their reading systems. Not surprisingly, it is patient E.M. who showed spared reading abilities after being diagnosed with SD.

Our implementation of the varying amounts of reading experience focused on the ratio between orthographic and visual input. One can think of it as the ratio of how much time an individual spends reading versus watching TV. Naturally,

such a statistic would correlate with other demographics such as years of education and occupation. An alternative would be to look at a network with more training versus less training on reading. We chose to focus on the ratio variable because of its correspondence to stable differences in the relative amount of time spent in activities such as reading and writing. Evaluating the impact of absolute rather than relative reading experience is left to future work.

Motivation for the reading pathway capacity manipulation

Just as reading proficiency heavily depends on how much time one spends reading, it also depends on the neural substrate used for reading. The capacity of that neural substrate may be shaped or altered by experience, but it is also biologically constrained, and there are individual differences along both of these dimensions. For example, developmental dyslexia is a condition characterized by underdeveloped reading skills (compared to age-matched controls) despite normal intelligence. Behaviourally, this can be seen in significantly slower and more error-prone word and pseudoword reading in dyslexic children. Functional neuroimaging studies have found that the posterior areas associated with orthographic processing and the integration of orthographic and phonological codes (i.e., occipito-temporal and temporo-parietal regions) consistently show decreased activation compared to those of nondyslexics in a range of lexical tasks (Pugh et al., 2001). This is the case in adult dyslexics even when they perform like the control subjects behaviourally (Brunswick, McCrory, Price, Frith, & Frith, 1999; McCrory, Mechelli, Frith, & Price, 2005). Thus, in early adulthood, even when the reading skills of dyslexics have improved sufficiently so that in nonspeeded reading conditions they achieve high levels of accuracy, the underlying neurobiological differences are still there and can be observed using functional neuroimaging (McCrory et al., 2005).

While the data on dyslexia reviewed above treat reading ability (and the underlying neurophysiological differences) dichotomously, is seems likely that these traits vary continuously, so that there are differences in the relevant pathways even within the population of individuals categorized as nondyslexics.

Motivation for the lesion distribution manipulation

Our final manipulation is motivated by the fact that no two patients have the exact same lesion. Even though it is well documented that in semantic dementia the atrophy starts from the temporal poles, affects predominantly the anterolateral temporal cortex, and progresses towards the posterior of the temporal lobes (Galton et al., 2001; Mummery et al., 2000, Whitwell, Anderson, Scahill, Rossor, & Fox, 2004), the rate of progression, the specific subregions affected, and the relative tissue loss in each of these regions may vary. So far, unfortunately, there has not been a detailed investigation of individual differences in the temporal lobe atrophy and its progression, and how that relates to performance on clinical tests.

One study looking at hemispheric differences in the lesion distribution in SD patients and how this relates to task impairment found that, even though in most SD cases the temporal atrophy is bilateral but more pronounced on the left, there is substantial variability (Lambon Ralph, McClelland, Patterson, Galton, & Hodges, 2001). There are cases where the atrophy is more pronounced in the right hemisphere. Importantly, it was also found that the degree and orientation of the asymmetry was correlated with the patients' relative performance on naming and word–picture matching.

A possible relationship between processing of words versus pictures and laterality has also been suggested by another study, which looked at SD patients' knowledge of famous people (Snowden, Thompson, & Neary, 2004). The laterality of the temporal lobe lesion was correlated with performance with names versus faces. Patients with predominantly right atrophy performed better with names while those with predominantly left

atrophy performed better with faces. The authors interpreted their findings as supporting the view that the anterior temporal lobes (ATL) are crucial for semantic processing, which involves the integration of information across different modalities, with the left ATL being particularly important for verbal processing and the right for visual (Snowden et al., 2004).

In addition, there have been a few reports about common regions of brain damage in groups of SD patients (Chan et al., 2001; Galton et al., 2001; Mummery et al., 2000; Rosen et al., 2002; Studholme et al., 2004). While findings agree on the anatomical regions most severely affected in SD, there is great variability in the full set of implicated regions reported by the different studies. If there are differences in the damaged brain regions reported for groups of SD patients, then there must be substantial differences between individuals taken from the different groups. Of course, this kind of reasoning applies also to reports of performance on clinical tests and the correlations between affected brain areas and performance.

While there is clearly variability in lesion distribution, there is very little information in any of these reports on the relation between lesion distribution and performance on reading and naming. Indeed, few of the studies use more than one task, and the two that do (Galton et al., 2001; Lambon Ralph et al., 2001) do not include reading. Clearly, this is an issue that requires further investigation. For now, it seems fair to conclude that there is a great deal of variability among the patients with respect to the distribution of their lesions. The third manipulation in the neural network model presented here explores the effects of such variability on naming and reading performance.

Using the three manipulations outlined above, we examined whether the model can fit the naming and reading data from five SD patients reported in the literature. The five patients selected for these simulations were chosen for two reasons: (a) to include patients who were all tested with the same materials, and (b) to include one of the three patients discussed above who does not show the usually observed association between semantic and lexical deficits—patient E.M. Three of the patients were reported in Graham, Hodges, and Patterson (1994) and were tested on a set of materials first used in that paper; the other two patients, including E.M., were tested on exactly the same materials by Blazeley et al. (2005). To anticipate, the results show that the data from all of the patients can be captured, and in all cases, the data are consistent with several different simulations involving one or more of the three manipulations. As suggested earlier, these findings have implications for the long-standing tradition in cognitive neuropsychology to view each case separately and to focus on individual reports of association and dissociation of symptoms. The results from our investigation support an approach whereby each patient is seen as a member of a population within which there is variability largely due to individual differences (and not simply noise).

Method

Patient data

The five patients included in the current investigations were the three patients initially reported in Graham et al. (1994), J.L., F.M., and G.C., and the two patients reported by Blazely et al. (2005), E.M. and P.C., who were tested with the same materials. The set includes 106 pictureable nouns, half of which have regular spelling-to-sound correspondences and the other half irregular. The two halves were matched for syllable length and frequency and formed three frequency bands. The advantage of using this set is that the patients can be tested on all tasks with the same materials, which allows for a detailed analysis of the error patterns across tasks.

Figure 2 shows the patient data broken down by frequency for naming and by frequency and regularity for reading. The original materials included three frequency bands but the simulation materials had only two bands (high frequency, HF, vs. low frequency, LF). Therefore, for the purposes of

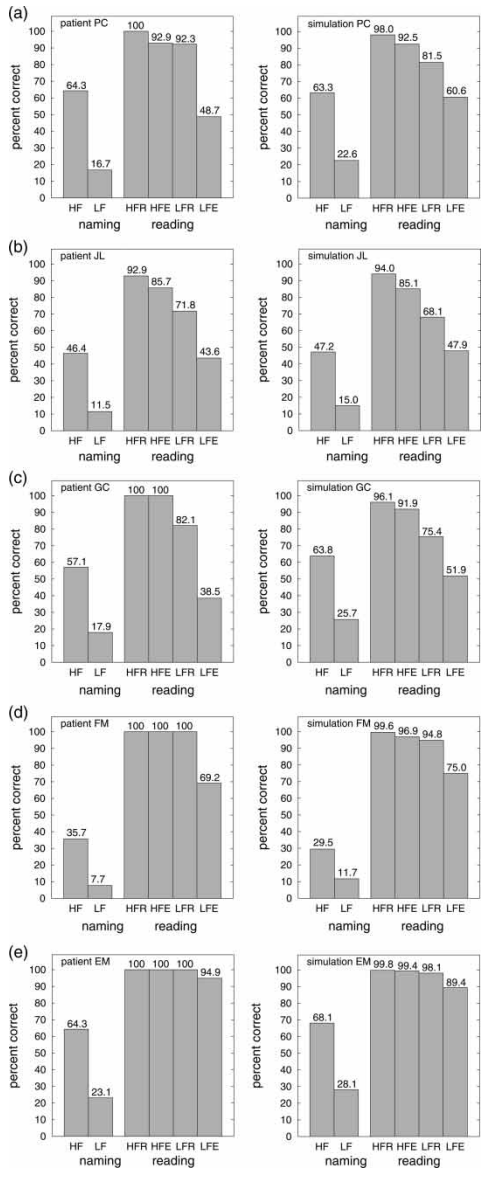

Figure 2. *Patient data along with representative successful simulations for each of the five patients: (a) P.C.: simulation fit with training O:V = 1:1 (O:V = orthographic-to-visual ratio), direct pathway of 20 units, and unbiased lesion; (b) J.L.: simulation fit with training O:V = 1:1, direct pathway of 20 units, and unbiased lesion; (c) G.C.: simulation fit with training O:V = 1:1, direct pathway of 20 units, and 75% orthographically biased lesion; (d) F.M.: simulation fit with training O:V = 2:1, direct pathway of 30 units, and 100% visually biased lesion; (e) E.M.: simulation fit with training O:V = 2:1, direct pathway of 20 units, and 100% visually biased lesion.*

the simulation, low- and medium-frequency items were combined in the results reported here.

Network

Network architecture. The architecture of the neural network is shown on Figure 3. It includes four input/output layers: orthography, phonology, vision, and action. There is full bidirectional connectivity between the input/output (also called *visible*) layers and the integrative hidden layer and full recurrence within the hidden layer. In addition, there is a fully recurrent direct-pathway hidden layer between orthography and phonology. Also, there are five task units that can be turned on or off to regulate which of the layers participate in a given task. This is implemented by having a very strong negative bias on all other units in the network, so that in the absence on input from the task units, the other units are virtually insensitive to inputs and do not participate in processing or learning. Activating a task unit raises the resting level of units in the participating layers up to −3.00, a value that then allows further excitatory input to bring the units into play during processing.

The inclusion of task units like those used here originates with Cohen, Dunbar, and McClelland (1990) and has subsequently been used in other networks (e.g., Plaut, 2002) where production of one of several alternative responses to a particular

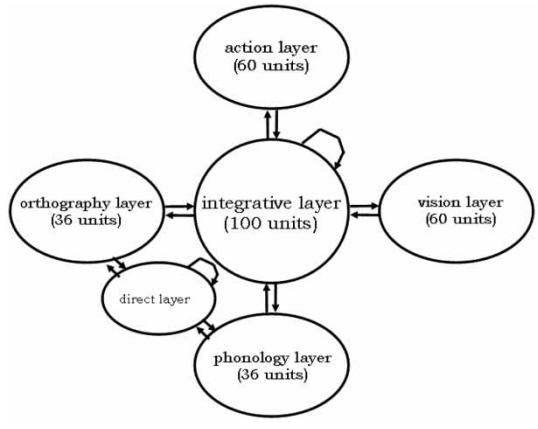

Figure 3. *Network architecture.*

input and/or responding based on one of several different available inputs is required. In our case, the use of task units also allows different strategies for performing a single task. For example, word reading could in principle be carried out by the direct pathway alone or by way of the integrative layer alone, or it could be carried out with both participating. The presence of task units during training encourages the network to develop the ability to do each task with only those parts of the network that are presently allowed to participate, instead of requiring the engagement of the entire network independent of the task.

Simulation materials. The same patterns were used for both training and testing. They consisted of 240 items from 12 categories (see Table 1). The visual patterns corresponded to visual representations of the object, and each of the 240 items had a unique visual representation. The action patterns, on the other hand, corresponded to representations of how one interacts with these objects, and each item did not necessarily have a unique action pattern. For example, we do more or less the same things with many types of fruit; peaches and nectarines are not really treated very differently. Both the visual and the action representations were 60-item long binary patterns that were generated randomly from a set of 12 category prototype patterns. Each prototype specifies the probability of occurrence of each of several binary values, so that items within categories are drawn essentially from the same probability distribution. These prototypes and the algorithm of pattern generation are shown in Appendix A. They were created using a procedure similar to that used in Rogers et al. (2004a). The prototypes and their similarity to each other were adjusted by hand to approximate the pattern of within- and between-category similarity of the patterns used by Rogers et al. (2004a).

Table 1. *Distribution of items in the semantic categories included in the simulation materials*

Category	*No. items*
Mammals	55
Reptiles	10
Birds	35
Fruits	20
Veggies	20
Musical instruments	15
Clothes	25
Utensils	10
Tools	15
Appliances	10
Vehicles	15
Furniture	10

The phonological and orthographic patterns were not real English words in general but were designed to approximate English spelling–sound consistency. They had a CVCC (where C is consonant, and V is vowel) structure with 12 possible onset and coda consonants (with matching graphemes and phonemes), 12 possible vowel graphemes, and 8 possible vowel phonemes. The only irregularities between spelling and pronunciation were in the vowels.

There were four groups of two vowel phonemes and three vowel graphemes as shown in Table 2 (60 items in each of these four groups). Every group comprised five types of items. Each of the five occupied cells was further divided into high-frequency and low-frequency items. The exact number of each type of item is shown in Table 3. These numbers are based on about 50,000 spoken word lemmas from the Celex English Lemma Database, where a lemma is defined as including all of the inflected forms of a word—for example, "dog" and "dogs" belong to the same lemma, and their frequencies of occurrence are added together to produce the lemma frequency (Burnage, 1990). The analysis of the corpus proceeded as follows: For each of the 20 grapheme–phoneme

Table 2. *Vowel phoneme–grapheme combinations*

	Grapheme		
Phoneme	*1*	*2*	*3*
1	regular	irregular	
2	irregular	regular	regular, rare

Table 3. *Number of high-frequency and low-frequency items in each of the four groups of vowel phoneme–grapheme combinations*

	Grapheme					
	1		*2*		*3*	
Phoneme	*HF*	*LF*	*HF*	*LF*	*HF*	*LF*
1	4	33	1	1		
2	1	1	2	12	1	4

Note: HF = high frequency. LF = low frequency.

correspondences (four groups of five types of items), the number of monosyllabic monomorphemic lemmas that matched that type of item was found. Frequency was also taken into account. High-frequency words were defined as having frequency of more than 70, while low-frequency words were defined as having frequency of less than 30 (Kucera & Francis, 1967). The relevant numbers were added to produce 10 sums—two (HF and LF) for each of the five types of items. These were then normalized to add up to 60—and these are the numbers presented in Table 3.

The phonological and orthographic prototypes and the method of pattern generation are presented in Appendix B. Once all the patterns were generated, the two groups—visual and action patterns on one hand and phonological and orthographic patterns on the other—were randomly matched to produce 240 items each with four patterns—visual, action, phonological, and orthographic.

Network training. Training consisted of a series of pattern presentations. In each presentation, the network was given either a visual or an orthographic pattern as input and was trained to produce either just the phonological or all four patterns as output. Processing in the network was regulated by a set of five task units. The five task units were used as summarized in Table 4. If no task unit is turned on, even in the presence of input, nothing happens in the network because all layers have a strong negative bias keeping the units' activations at 0. When a task unit is turned on, it sends positive activation to the layers it is connected to, effectively eliminating the negative bias of those layers and encouraging them to participate in the task at hand. Table 4 lists all the task units and the layers they are connected to. Task units are not uncommon in connectionist networks. They were included in order to encourage the network to make full use of all the pathways. For simplicity, we hard-coded the connection weights between the task units and the layers; other models have used training to learn task weights (e.g., Plaut, 2002).

The relative occurrence of visual versus orthographic input was one of the manipulations and is explained further in the experimental design section. However, for each of these two kinds of input, the relative occurrence of the requested output (phonology vs. all four output patterns) had the constant ratio of 1:1. Furthermore, when the network was trained to read (i.e., given the orthographic input to produce the phonological output), a third of the time only the direct pathway was used, while the rest of the time both the direct pathway and the integrative layer participated.[1]

A frequency manipulation was applied to both visual and orthographic training so that high-frequency items were seen 8 times more often than low-frequency items. Also, the different training tasks were not blocked. The network was trained on all items and tasks in an interleaved manner, and the order of the items was random. Back-propagation was used to update weights between units after every example. The presentation of each example lasted for seven simulated unit time intervals, each divided into a number of ticks; in each tick, net inputs to units were adjusted according to:

$$\Delta n_i = \frac{1}{(\text{number of ticks})} \sum_j (a_j w_{ij} - n_i),$$

[1] Because the majority of the learning involves semantic connections, those connections quickly grow in size and thus become responsible for the error in subsequent learning. The direct pathway needed to be trained by itself in order for any learning to occur for its connections.

Table 4. *Training details*

Input	*Output*	*Relative occurrence*	*Task unit/task*	*Participating layers*
Visual	Phonology	1/2	Name	Visual, phonological, integrative
	All	1/2	Think (from vis input)	All but direct pathway
Orthography	Phonology	1/3	Read for meaning	Orthographic, phonological, integrative, direct pathway
		1/6	Read	Orthographic, phonological, direct pathway
	All	1/2	Think (from orth input)	All

Note: vis = visual. orth = orthographic.

where j indexes the units connected to the current unit i, a_i and n_i are, respectively, the activation and the net input to unit i, and w_{ij} is the value of the connection weight to unit i from unit j. After each unit's net input is updated, the activation is also updated, based on the logistic function:

$$a_i = \frac{1}{1 + e^{-n_i}}.$$

The number of ticks determines how closely the network approximates the assumed underlying continuous evolution of activations. We used 4 ticks per interval, which provides a reasonably smooth evolution of the activations of the units in the network (see Plaut et al., 1996).

The seven time intervals were subdivided as follows. During the first three intervals, the visual or orthographic input pattern corresponding to the item being processed was clamped onto the appropriate layer. For the remaining four intervals, the input was removed, and the network was allowed to adjust the activation of all units in all layers, including the one previously clamped. During the final two intervals, the activations of units are compared to their corresponding targets (which were patterns over the phonological layer only or over all four visible layers, as described earlier). The relevant task unit was clamped on for the entire duration of the example presentation.

The network was trained using standard gradient descent with no momentum. The learning rate was set to 0.001 and the weight decay to 0.000001. Training continued for 2,500 sweeps through the training set.

Network testing. During testing, the network was again presented with either the orthographic or the visual pattern of each item. The phonological response was determined by selecting the most active units at each of the onset, vowel, and coda positions. Responses were either correct or incorrect depending on whether the network was able to exactly produce the actual phonological pattern. For both tasks, all layers were encouraged to participate during testing (by using the *think from orthographic input* task unit). At the end of training, performance on both reading and naming was perfect.

The network was tested following damage by selectively removing units in the integrative hidden layer as well as connections between that layer and the four visible layers. The damage to connections could be unbiased, visually biased, or orthographically biased, and the degree of bias could be 50%, 75%, or 100%. In all cases, when x% of integrative units were removed, x% of incoming and x% of outgoing integrative connections were also removed, where x is an average over the four sets of bidirectional pathways between the visible and the integrative layer. In the unbiased lesion, x% of connections were removed between the integrative layer and each of the visible layers. In the 50% biased lesion, 50% of the damage was in the direction of the bias. That is, $2x$% of the connections in the direction of the bias (visual or orthographic) were removed. The remaining damage was equally distributed among the remaining three pathways, so that $0.67x$% of the connections were damaged in each of those three pathways. In the 75% biased lesion,

$3x$% of the connections in the direction of the bias were removed, and $0.33x$% of the connections were removed in the other three pathways. Finally, in the 100% biased lesion, $4x$% of the connections were removed in the directions of the bias, and all other connections were unaffected. Ten levels of damage were examined (1–10%). To ensure appropriate sampling, the model was tested 20 times for each combination of lesion extent and distribution (using 20 different random number generator seeds).

Experimental design

The current project measured the relationship between naming and reading performance of the network. Reading was investigated as a function of the severity of the naming deficit and the values of three parameters: (a) training regime, (b) direct pathway size, and (c) lesion distribution. In order to map out the form of this relationship, the network was tested at 10 levels of lesioning, as mentioned above. A baseline was chosen for each parameter: The baseline training regime included visual and orthographic input in the ratio 1:1; the baseline direct pathway was 20 units, and the baseline lesion distribution was unbiased. The baseline level simply represents a point of comparison, chosen to be neutral and/or intermediate between the other values of the parameters explored. Each parameter was then manipulated individually, so as to promote either better or worse reading performance than that at baseline.

The training regime had an orthographic-to-visual input ratio (V:O) of 1:1 (baseline), 1:2, or 2:1. Having more experience with the orthographic labels than baseline (i.e., the V:O = 2:1 condition) supported better reading than baseline, while having less experience with the labels than baseline (i.e., the V:O = 1:2 condition) supported worse reading than baseline. The direct pathway size was 10, 20 (baseline), or 30 units. Having a larger or a smaller direct pathway fostered respectively better or worse reading than baseline. Finally, the lesion distribution was unbiased (baseline), visually biased, or orthographically biased. As explained earlier, the bias of the lesion had three degrees: 50%, 75%, or 100%. The 100% visually biased and the 100% orthographically biased lesions were the most extreme manipulations of this parameter. Visually biased lesions supported better reading than the baseline unbiased lesion, while orthographically biased lesions were worse at reading than baseline.

In summary, this was a $3 \times 3 \times 7$ full factorial design. All combinations of the three parameters were allowed.

Results and discussion

Fitting the patient data

In order to evaluate the ability of the model to fit the five patients' data discussed earlier, the following steps were taken. For each combination of values of the three network parameters, we first selected an appropriate lesion extent to best match each patient's overall naming performance. We then assessed the network's performance on both naming and reading at that lesion level in relation to the patient's data. The criterion for a successful fit was that for each task and item type (HF vs. LF for naming, and high-frequency regular, HFR, vs. high-frequency exception, HFE, vs. low-frequency regular, LFR, vs. low-frequency exception, LFE, for reading) the patient's data fitted within the 95% confidence interval of the network.

That is, we treated the network's proportion correct on each item type in each task as if it represented the underlying probability of correct performance in that condition for the patient.[2] We then calculated whether the observed patient proportion correct fell within the 95% confidence

[2] Note that the network values were based on 240 items each tested 20 times; while there may be some uncertainty in these values due to the fact that they are based on a random sampling process, the number of samples is such that the variability is small enough to have a negligible effect in these analyses.

interval of that value—that is:

$$net \pm 2\sqrt{\frac{net(1-net)}{n}},$$

where *net* is the network's proportion correct, and *n* is the actual number of items on which the patients' performance had been tested. Finally, we used a chi-square test with four degrees of freedom to confirm that the reading performance of each simulation was indeed a good fit for the patient data. The three network manipulations of training regime, direct pathway size, and lesion bias were applied individually and in combination for each of the patients.

With the neutral baseline parameters (training regime of O:V = 1:1, direct pathway of 20 units, and an unbiased lesion), successful fits were found for two of the patients: P.C. ($\chi^2 = 5.494$, $p = .240$) and J.L. ($\chi^2 = .531$, $p = .970$). There were many more fits for both of these patients when the three network parameters were manipulated, individually or in combination. None of the other three patients was successfully modelled by the baseline simulation. However, all of them—including patient E.M., who was the dissociation case reported by Blazely et al. (2005)—also had multiple fits when the parameter values were varied. Appendix C includes a full list of the successful fits for all five patients. Figure 2 shows the performance predicted by one of the best fitting simulations for each patient, along with the patients' actual data. The simulations shown for P.C. and J.L. are networks with baseline parameters, the simulation for G.C. is a network with baseline training, a direct pathway of 20 units, and 75% orthographically biased lesion (fit: $\chi^2 = 5.930$, $p = .204$), the one for F.M. has a training regime O:V = 2:1, a direct pathway of 30 units, and 100% visually biased lesion (fit: $\chi^2 = 2.646$, $p = .619$), and finally, the one for E.M. has training O:V = 2:1, a direct pathway of 20 units, and 100% visually-biased lesion (fit: $\chi^2 = 2.141$, $p = .710$).

As can be seen in Appendix C, patient J.L. had the greatest number of successful fits. Many different combinations of the three parameters produced a pattern similar to J.L.'s data. Generally, most fits were from networks with a direct pathway of 20, a baseline training of 1:1, or training O:V = 2:1, and an unbiased or slightly biased lesion (in either direction). This pattern confirms that the baseline simulation was indeed most suitable for fitting J.L.'s data. A similar trend is seen for P.C.'s fits—they tend to involve a baseline training regime and direct pathway size and an unbiased or slightly biased lesion. Patient G.C.'s fits also involve baseline training and direct pathway, but predominantly lesions with orthographic bias.

Turning to the two more extreme cases, F.M. and E.M., there was a clear preference for a training of O:V = 2:1, a larger direct pathway, and a visually biased lesion. Interestingly, there were more fits for E.M. than for F.M., and the fitting tendencies for all three parameters were more marked for F.M. than E.M. This observation belies the idea that patient E.M. has any "special" status as a dissociation case. We were able to fit all five patients with multiple simulations manipulating the same three network parameters. If the four patients who exhibited surface dyslexia come from a single distribution then E.M. falls within that distribution as well and is no more of an outlier than patient F.M.

It may be, in fact, that it is patient F.M., rather than patient E.M., who is the outlier in this series. Unlike the other patients in this set, F.M. showed an extreme deficit in naming relative to her performance on other tests of semantic knowledge including word–picture matching. An explanation for this (consistent with the earlier analysis of this patient by Lambon Ralph et al., 2001) is that in the case of F.M. there is a special problem in the connections from semantics to phonology. Further consideration of this possibility is beyond the scope of the present simulations and is left for future research.

Our simulation involved three individual differences variables. Are all three necessary to account for the individual patient data? We now consider each of the three individual differences variables in turn. Considering first the lesion distribution

variable, it appears that this factor alone is enough to allow an account for the data of all five of the patients. That is, for each patient, there is at least one simulation that falls within the 95% confidence interval of the data while the other two individual differences variables (direct pathway size and training regime) remain at baseline levels. Also of considerable interest is the fact that four of the five patients can be fitted by manipulating only the training regime variable: J.L., P.C., G.C., and E.M. all can be fitted with baseline values of the direct pathway and lesion distribution variables. It is thus possible to account for the data from patient E.M. by assuming only an experience manipulation (although in this case the fit is not as good as others involving a combination of training regime and other variables). Varying the direct pathway size by itself allowed the model to account only for the three patients J.L., P.C., and G.C., although it is possible that a more extreme direct pathway manipulation would have allowed a fit to E.M. and F.M.

Relationship between reading and naming

A multiple regression was performed to look at the relative contribution of each of the three individual difference parameters to the relationship between naming performance and reading of irregular items. For each combination of the three parameters, a set of naming and reading data was obtained at the 10 lesion levels. The analysis included one outcome—the logit of the reading performance on all irregular items—and four predictors: training regime, direct pathway size, lesion bias, and the logit of the naming performance.[3] The training regime was quantified by coding the baseline (O:V = 1:1) as 0, less orthographic training (O:V = 1:2) as −1, and more orthographic training as 1. Similarly, the lesion bias was coded as 0 for unbiased lesions, −100, −66.7, and −33.3 for orthographically biased lesions (most extreme to least extreme), and 100, 66.7, and 33.3 for visually biased lesions. This coding preserved the linear relationship among the different biases as implemented in the network where the bias levels are 100%, 75%, and 50%, and the unbiased lesion is in fact 25% in each direction.

Not surprisingly, naming impairment and reading impairment were found to be highly correlated, $R^2 = .946$, $t(58) = 4.76$, $p < .0005$. More importantly, each of three manipulations were significant predictors of the reading deficit after controlling for naming impairment as well as the variance accounted for by the other manipulations: training regime manipulation, $R^2 = .609$, $t(58) = 11.10$, $p < .0005$; direct pathway size manipulation, $R^2 = .140$, $t(58) = 15.48$, $p < .0005$; lesion bias manipulation, $R^2 = .940$, $t(58) = 7.72$, $p < .0005$. The lesion bias was found to account for the greatest amount of unshared variance.

Figure 4 illustrates the spread produced by the three manipulations in the relationship between naming impairment and reading of irregular items. While lesion bias resulted in the largest spread compared to the other manipulations when naming was relatively low (<30% correct; which is where the majority of the data points are), it is the training regime manipulation that produced the largest spread at higher levels of performance and also produced a relatively high spread at lower levels of performance. Finally, the direct pathway manipulation consistently resulted in the lowest spread at all levels of performance. For example, at naming 20% correct, the difference in reading performance between the two extremes of the direct pathway size was 23%, the difference between the two extremes of the training regime was 29%, and the difference between the two extremes of the lesion bias was 31%; on the other hand, at naming 50% correct, this difference was 9% for the direct pathway manipulation, 10% for the lesion bias manipulation, and 18% for the training regime manipulation. Our model, therefore, has an interesting implication: While overall the distribution of the

[3] $\text{logit}(x) = \log[x/(1 - x)]$. The raw probability correct was replaced by the respective log odds of success in order to avoid effects of logistic compression in the performance.

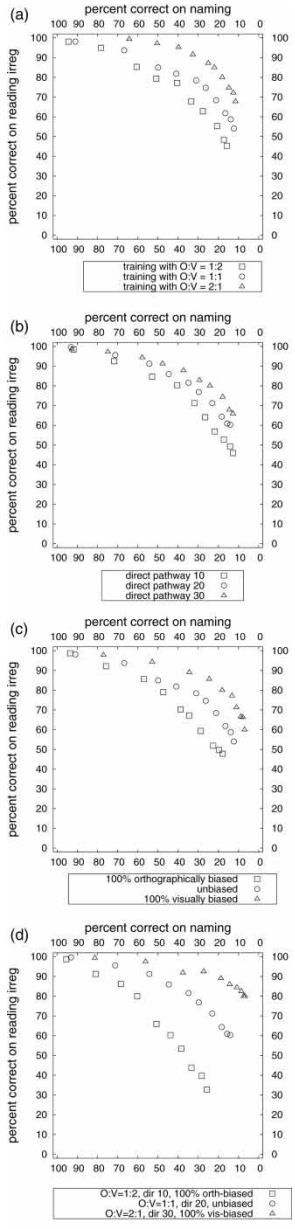

Figure 4. *The effect of the three manipulations on the relationship between naming deficit and impairment on irregular-word reading in the damaged model: (a) effect of training regime (with a direct pathway of 20 and unbiased lesion); (b) effect of direct pathway size (with training of O:V = 1:1 and unbiased lesion; O:V = orthographic-to-visual ratio); (c) effect of lesion bias (with training of O:V = 1:1 and direct pathway of 20 units); (d) cumulative effect of the three factors (only the most extreme cases shown).*

lesion has the most profound impact on how well naming and reading performance correlate in a given patient, it is perhaps individual differences in premorbid reading experience that can explain variability in the correlation between naming and reading performance early in the course of the disease.

Characterization of network behaviour

Thus far we have analysed our networks by showing how they exhibit sensitivity to particular factors and examining goodness of fit to individual patients. Here we consider how the network's behaviour might be captured by a reduced mathematical description, for the purposes of summarizing in a succinct way how its performance is affected by different variables. This characterization relies on the idea that the underlying strength of the network's tendency to accurately perform a particular task may reflect a variety of factors in a simple, perhaps additive way, while the relationship between this strength variable and overt performance has a sigmoidal or logistic shape. With such a function, performance reaches a ceiling level as strength increases such that further strengthening will have little effect, and similarly, below a certain level performance is at floor so that further weakening will have little effect. We pursued this idea by asking how well we can capture the effects of our manipulations of the network's performance in a logistic regression, where lesion extent, direct pathway size, training regime, lesion bias, frequency, and regularity are all factors that should contribute to the strength of the tendency to read an item correctly. The idea is based in part on an analysis previously presented in Plaut et al. (1996), indicating how frequency and regularity both contribute in an additive way to the strength of the connections subserving a particular grapheme–phoneme correspondence in a highly simplified version of their model of single-word reading.

Here, we performed a separate logistic regression of the reading data for each level of each of the three network parameters—that is, direct pathway size, training regime, and lesion bias (while the other two parameters were held

constant at baseline), as well as for the two most extreme cases of combining the three manipulations. Four factors were entered as predictors in these analyses—the lesion extent, the frequency and regularity of the items, and an interaction term coding for the frequency-by-regularity relationship.

One of the most important findings was that the frequency-by-regularity interaction was not significant in any but one of the nine logistic regressions (for the 100% orthographically biased lesion $p = .033$; all other $p > .05$), while lesion extent, frequency, and regularity were all highly significant (all $p < .0005$). The occurrence of only one significant frequency-by-regularity interaction term out of nine logistic regression tests suggests that the frequency-by-regularity interaction found in an analysis of variance (ANOVA) run on the raw performance data, $F(1, 171) = 298.78$, $p < .0005$; see Figure 5) and seen in the patient data (e.g., Patterson et al., 2006) is simply a consequence of the compression of performance near the high end of the logistic function. Using the coefficients given by the logistic regression analyses, we plotted the relationship between lesion extent and reading performance. On the same graphs we also plotted the actual data from the network. Three of these graphs, showing the baseline case and the two most extreme cases of combining the three manipulations, can be seen in Figure 6. The logistic functions generally provided excellent fits to the data. One slight exception is that the network performance in the case of the largest

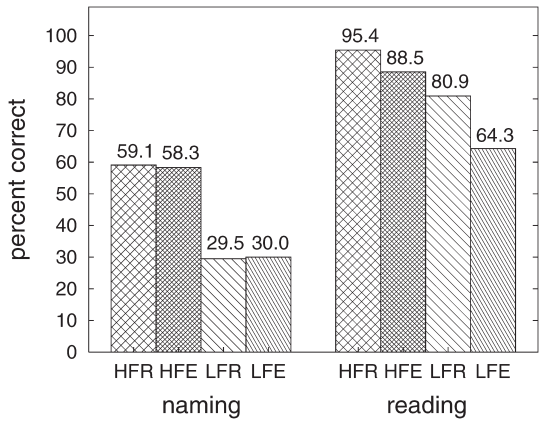

Figure 5. *Frequency by regularity by task three-way interaction in the network. HFR = high-frequency regular. HFE = high-frequency exception. LFR = low-frequency regular. LFE = low-frequency exception.*

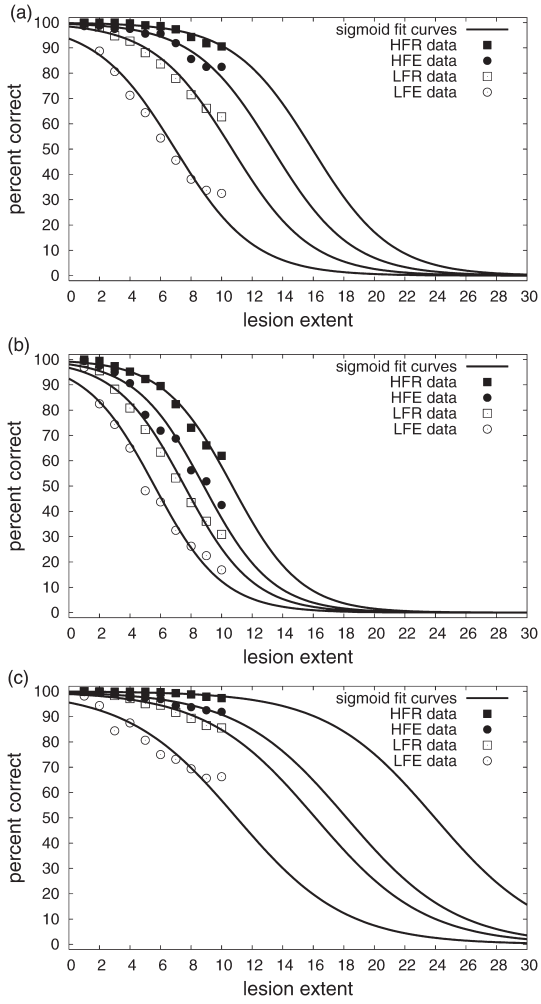

Figure 6. *Logistic regression of the reading data: Effects of frequency and regularity: (a) in the baseline simulation: direct pathway of 20, training regime O:V = 1:1 (O:V = orthographic-to-visual ratio), and unbiased lesion; (b) in simulation with direct pathway of 10, training regime O:V = 1:2, and 100% orthographically-biased lesion; (c) in simulation with direct pathway of 30, training regime O:V = 2:1, and 100% visually biased lesion.*

lesions seems to be consistently better than the prediction of the logistic function, which suggests that the effect of lesioning may be levelling off, reflecting the residual functional capabilities of the direct pathway.

These figures illustrate that reading performance falls off with lesion extent in a sigmoid-like fashion, and frequency and regularity are two independent parameters that "shift" the position of the sigmoid curve to the left or the right (see also Plaut et al., 1996). Lower frequency items have a fall-off curve more to the left than higher frequency items, which is why they are impaired earlier and to a greater extent. Similarly, irregular items have a fall-off curve more to the left than regular items. When the two effects are superimposed on each other, they result in what appears as an interaction—the low-frequency irregular items are significantly more impaired than the other three types of item.

Next we consider how the reading performance fall-off curve is affected by the three network manipulations. There are two possibilities—an effect on the intercept and/or an effect on the slope. An effect on the intercept is analogous to the frequency and regularity effects described in the previous paragraph. An effect on the slope, on the other hand, is a change in the curve fall-off rate, rather than a change in its position. Figure 7 illustrates the relationships between the three factors and the intercept and slope of the logistic curve. It can be seen that the direct pathway size manipulation affected the intercept most strongly of all three manipulations and did not affect the slope. Similarly, the training regime manipulation affected the intercept and hardly at all the slope. In contrast, the lesion bias manipulation had a strong effect on the slope and no reliable effect on the intercept. This last effect was not surprising given the implementation of the bias manipulation—for any one lesion extent, as we go from an orthographically biased lesion through unbiased to visually biased, the amount of severed links which contribute to reading decreases, which is equivalent to larger versus smaller lesion with respect to reading performance.

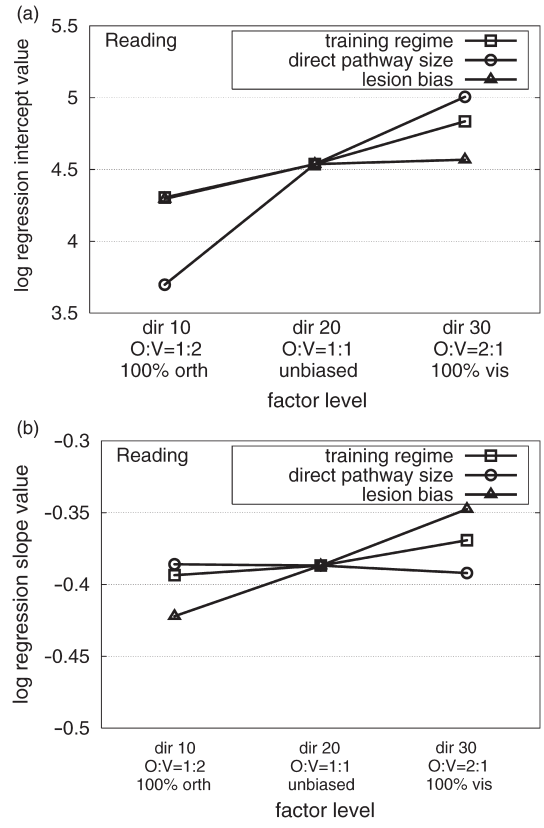

Figure 7. *Logistic regression of the reading data: Effects of the three network manipulations on the intercept and the slope of the logistic function: (a) effects on the intercept; (b) effects on the slope. O:V = orthographic-to-visual ratio.*

The effects described above suggest that premorbid factors such as the amount of reading experience and the capacity of the neural pathway mapping from orthography to phonology affect the susceptibility of reading performance to damage. More experience and better orthographic-to-phonological mapping make reading performance more robust by delaying the point at which performance starts to fall off. On the other hand, postmorbid factors such as the extent and distribution bias of the lesion affect the rate at which reading performance falls off with damage. A lesion oriented away from the orthographic input layer and more towards the visual object input has less effect on reading, as evidenced

by the decreased rate at which performance falls off.

It is worth noting that one reason why training only slightly affected the rate of decrease in reading ability may be the fact that there was no retraining after initial lesioning. In contrast, semantic dementia is a progressive disease, and the patients have a chance to try to maintain their skill even as their condition worsens. Using a feed-forward connectionist network, Welbourne and Lambon Ralph (2005) showed that continued training while a network underwent a progressive loss of connections led to considerable preservation of reading ability. Reading habits and occupation, therefore, may relate not only to the initial susceptibility of reading performance to damage, as we have seen here, but also to the rate at which this performance falls off.

Turning to the other task, naming, we performed a similar set of analyses. Initially, we included the same four predictors as those for reading: lesion extent, frequency, regularity, and frequency by regularity. However, regularity and the frequency-by-regularity interaction were not significant in all but one of the nine logistic regressions (for the network combining training O:V = 1:2 with direct pathway of 10, and 100% orthographically biased lesion $p < .05$; all other $p > .08$), while lesion extent and frequency were highly significant (all $p < .0005$). These findings confirmed the results of the ANOVA run on the raw performance data, which also indicated that frequency but not regularity affected naming performance (see Figure 5). This was expected since the regularity of the spelling-to-sound correspondence of a word is irrelevant for naming, where a spoken word is produced in response to a picture. Parallel to the network's performance, the patient naming data lack an effect of regularity and show a pronounced frequency effect. We therefore reran the analyses including only lesion extent and frequency as predictors. These analyses again confirmed the highly significant effect of lesion extent and frequency on naming (all $p < .0005$). Furthermore, they allowed us to explore how the naming performance fall-off curve was affected by the three network manipulations. The results are presented in Figure 8.

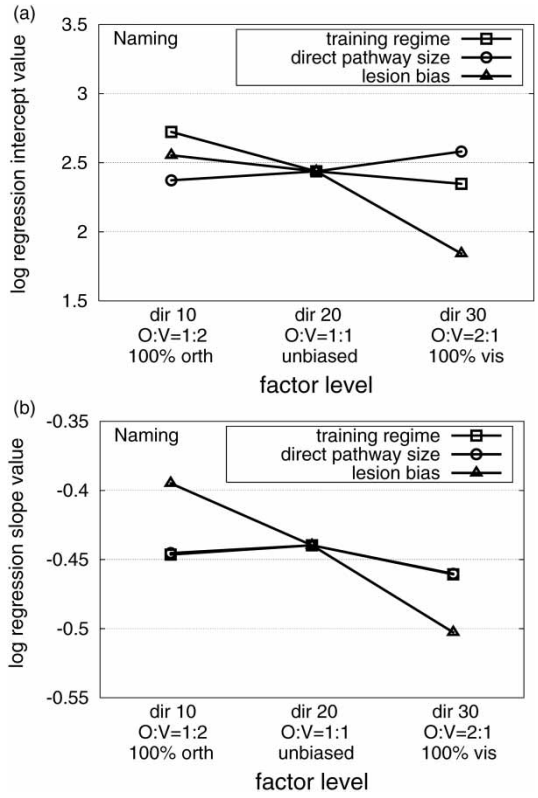

Figure 8. *Logistic regression of the naming data: Effects of the three network manipulations on the intercept and the slope of the logistic function: (a) effects on the intercept; (b) effects on the slope. O:V = orthographic-to-visual ratio.*

The slope of the naming fall-off curve was consistently steeper than that of the reading fall-off curve while the intercept was consistently smaller, indicating that naming was more sensitive to semantic lesions than was reading. There are at least two explanations for this. First, while reading involves a *systematic* mapping between graphemes and phonemes, naming involves an *arbitrary* mapping between the visual characteristics of an object and its name. This is true even though word reading is acquired later and with greater effort than naming; ultimately, reading is an "easier" task in that it involves a highly systematic mapping from one modality to another and is,

therefore, less susceptible to damage than is naming. Secondly, there is a direct pathway linking orthography to phonology, which is not affected by the lesion. Hence, the system has that pathway to rely on for reading but not for naming. Of course this pathway is usually not perfectly proficient in reading; that is, in the absence of semantics, there is some reading impairment. The extent of this impairment is a function of the capacity of the pathway and the specific reading experience that the network has had. Further simulations are required to investigate how a division of labour between the direct pathway and the semantic pathway in reading may arise during training as a function of pathway capacity and training regime (see also Harm & Seidenberg, 2004).

With respect to the three network manipulations, the analyses indicated that the direct pathway size had no effect on naming, which was expected since the direct pathway is not relevant for this task. In contrast, both the training regime and the lesion bias affected the naming curve. Similarly to reading, the training regime affected only the intercept of the logistic function and not the slope. The regression results indicate that similarly to reading, even though perhaps to a lesser extent, increased experience delays the point at which naming performance begins to decrease. Finally, the lesion bias had a very strong effect on the slope of the naming curve as well as a considerable effect on the intercept. This is different from the trend seen for reading, where the lesion bias only affected the slope. It suggests that the location of the lesion is especially important for naming; it is responsible for both the initial point of drop in performance and the rate of this drop as the disease progresses. A lesion oriented away from the orthographic input layer and more towards the visual object input has a great impact on naming.

To further investigate how lesion bias affects reading and naming performance, we conducted a simulation where the network was damaged at one location only. Three lesion locations were considered: units in the integrative layer; links between the integrative layer and the visual layer; and links between the integrative layer and the orthographic layer. In order to get comparable results to those obtained in the investigations of effects of lesions reported above,[4] in this final simulation, the lesions progressed over 25 levels, where level x represents a lesion of x% for integrative units and $4x$% for connections between the integrative layer and a visible layer. Thus, the level 25 of lesioning links was a complete obliteration of the connections between the relevant visible layer and the integrative layer.

The results can be seen in Figure 9. This simulation supported the idea that lesioning connections between semantics and the visual layer profoundly impairs naming performance and hardly affects reading performance, while lesioning connections between semantics and the orthographic layer results in substantial reading impairment and very little naming impairment. Finally, lesioning semantic units affects performance on both tasks but naming more so than reading. Notably, these results illustrate that, in the model, performance on each of the two tasks relies most heavily on certain pathways (visual object representation through semantics to phonology for naming and orthographic word representation through semantics and the direct layer to phonology for reading) but it is also influenced by the activation of other available information. This is an essential characteristic of a single, highly interactive system, which integrates information from different modalities and of different types and participates in a wide range of tasks encompassing the semantic and lexical domains. This is the role attributed to the anterior temporal cortex in our account of semantic dementia (Rogers et al., 2006; Rogers et al., 2004a).

[4] As explained earlier, in the original simulations, a 5% unbiased lesion involved lesioning 5% of the integrative units plus 5% of the incoming and 5% of the outgoing connections between the integrative layer and each of the four visible layers (adding up to a total of 20% damaged links); an equivalent lesion fully biased towards a specific layer involved lesioning 5% of the integrative units plus 20% of the incoming and 20% of the outgoing connections between the integrative layer and the specified visible layer (and no links between the integrative layer and any of the other visible layers).

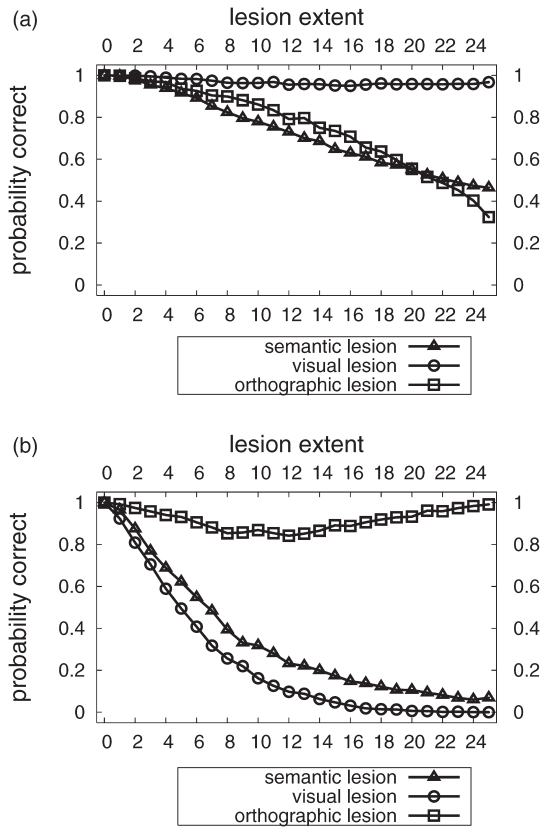

Figure 9. *Further investigation of the effect of lesion location on reading and naming: (a) effects on reading; (b) effects on naming.*

GENERAL DISCUSSION

In summary, the present computational investigation of the relationship between naming and reading deficits in SD adopted a single-system perspective using a connectionist model that implements semantics. In this model, information from different modalities and types feeds into a single integrative layer. The results from our simulations showed that within such a system, we can replicate the overall relationship between reading and naming seen in SD patients (as reported by Patterson et al., 2006, and Woollams et al., 2007), and we can successfully model data from five SD patients among whom there are cases of both association and dissociation of deficits. The driving force of the account was the manipulation of three network parameters seen as implementing plausible individual differences among the patients. The three factors were experience with reading, capacity of the direct pathway mapping orthography to phonology, and spatial distribution bias of the lesion. Each of these factors contributed significantly and uniquely to the variability in the relationship between naming and reading impairment.

We also captured the role of frequency and regularity in reading. Specifically, we showed that the frequency-by-regularity interaction seen in both the network's and the patients' reading performance can be accounted for by assuming that both frequency and regularity have additive influence on the underlying strength of the correct response to an item, but that this effect is subject to logistic compression.

There are three main points that emerged from our work. First, the pattern of association and dissociation between naming and reading seen in semantic dementia is consistent with our single-system account. The strong relationship between semantic and lexical deficits seen in semantic dementia patients is also found in our computational model, where both reading and naming depend on an intact semantic system; extremely rare patients like E.M. who appear to evidence dissociation between conceptual knowledge and lexical knowledge can, in fact, be accounted for in our single-system model. Second, premorbid and postmorbid individual differences among SD patients are likely to be important in accounting for the variability in patients' performance on semantic and linguistic tasks such as naming and reading. Third, there is a strong need for a thorough investigation of individual differences in temporal lobe atrophy in semantic dementia and how it relates to behavioural performance on semantic and linguistic tasks. These points are reviewed in the following paragraphs.

A single system for conceptual and lexical knowledge

Our simulations show how both an association and a dissociation of function can be produced in a

single system. Despite the high correlation between naming and reading of irregular words in our network ($R^2 = .95$)—which is indeed the trend found also in group studies of SD patients (cf. Patterson et al., 2006; Woollams et al., 2007)—we found multiple data fits for patient E.M. who shows little impairment in reading despite her profound naming deficit. While our results do not rule out a two-system perspective, they clearly show that the data from patients like E.M. is not, as others have argued (e.g., Blazely et al., 2005), inconsistent with the view that there is a single, cross-modal, highly interactive system that integrates different types of information (including linguistic information) and participates in a wide range of tasks encompassing the semantic and lexical domains.

Our results also underscore the danger of overfitting due to attributing too much weight to individual patient data. Rather, we need to be able to characterize the distribution and understand the variability in that distribution as a function of individual differences (see also Woollams et al., 2007). If there is in fact an underlying continuum as suggested here and in previous reports, it could be misleading to look at selected few individuals, for example P.C. versus E.M. reported by Blazely et al. (2005).

The importance of individual differences

The present computational study highlights the importance of individual differences in explaining neuropsychological data. Each of the three individual differences factors investigated—reading experience, capacity of the direct pathway mapping orthography to phonology, and bias of the lesion distribution—affected the impairment of reading and/or naming produced by damage, as well as the relationship between these two impairments.

In addition to the variability in overall naming and reading performance in the patient data discussed here, there is also a lot of variability in the reading data when broken down by item type. As mentioned earlier, the general trend in both SD patients and the network is such that performance on HF words is better than that on LF words, performance on words with a regular spelling-to-sound correspondence is better than that on irregular words, and the two effects are not independent so that LF irregulars are particularly vulnerable under damage. The extent of this interaction, however, varies greatly across patients. Interestingly, in our simulations, all three manipulations influenced this interaction, suggesting that individual differences may explain not only variability in overall reading impairment among patients, even after controlling for the semantic impairment as indexed by their naming performance, but also variability in the frequency-by-regularity interaction seen in reading. As discussed earlier, our logistic analyses showed that some of these effects may be additive, so that varying amounts of reading experience and direct pathway capacity, for example, may simply shift the reading curves of the different types of items (HFR, HFE, LFR, and LFE). We found that premorbid factors preferentially affect the susceptibility of reading and naming performance to damage, while postmorbid factors mainly contribute to the rate at which this performance worsens with damage. In addition, the lesion location is especially important for naming as it has considerable impact for the initial point at which performance starts to decline as well as the decline rate.

Individual differences in affected temporal subregions in semantic dementia?

Finally, the current investigation of the role of individual differences in the relationship between naming and reading impairment under semantic damage has suggested that lesion location may be an important factor affecting the exact relationship between performance on semantic and lexical tasks. This is a notion also embraced by the separate systems account (Blazely et al., 2005), which posits that the observed high correlation between semantic and lexical impairment in the majority of SD patients is explained by damage to two anatomically neighbouring but functionally distinct systems. Here we have shown that the specifics

of the brain damage can be very relevant to a single systems account as well. Therefore, a thorough investigation looking at individual patients' brain scans and relating region-specific atrophy to task performance would be highly informative to both accounts and possibly relevant to settling the debate between these two opposing views. Combined with demographic data about each patient—including relevant parameters such as years of education, occupation, reading habits, previous assessments of literacy, and so on—this kind of database may be able to address some of the predictions made by our model. One prediction is that premorbid reading experience may explain variability among patients early in the course of the disease, while lesion distribution may explain later variability as the disease progresses.

It may be worth noting, however, that our results do not strongly require an appeal to differences in lesion location to explain most of the available data. In particular, we were able to fit the data from patient E.M. with only a manipulation of the reading experience variable (O:V = 2:1), or with a combination of reading experience and direct pathway size (O:V = 2:1, direct pathway of 30). Among the patients considered here, only patient F.M. could not be fitted without manipulating the lesion distribution variable. This fact, together with the recent success of Woollams et al. (2007) in accounting for the data from a large number of SD cases with a single premorbid individual difference variable, suggests that it not yet clear exactly how important the lesion distribution factor is in explaining variability in the relation of reading and naming performance in semantic dementia.

Limitations of the present investigations

As a final note, we briefly mention several limitations of the present work. First, it has not considered lexical decision (LD), a task used with E.M. and P.C. in Blazeley et al. (2005), but not with the other patients tested. Blazely et al. found that LD was relatively preserved in patient E.M., consistent with their view that there is a separate orthographic lexicon, which is spared in this patient. We acknowledge that an account for patient E.M.'s data from the perspective of our single system account will not be complete until patterns of lexical decision as well as word reading are simulated.

More generally, the present investigation has not considered a wide range of other tasks that have been used with semantic dementia patients. Data from such tasks may further constrain our single-system model, and future research investigating a wide range of patients on a fuller battery of tasks will be important for further progress in understanding the relationship between lexical and semantic processes, in semantic dementia and in general.

CONCLUSION

Our results support the notion that semantic and lexical tasks do not require separate systems; the range of data patterns produced by SD patients is consistent with the idea that both types of tasks can be performed (and can deteriorate at varying rates upon damage) within a single system. Cases of association and dissociation of semantic and lexical deficits may well represent a combination of graded influences rather than any sort of underlying dichotomy. The results are consistent with the view that all SD patients come from a single distribution, the variance of which is a function of individual differences existing prior to brain damage and individual differences in the extent and spatial distribution of the progressive brain damage. Our investigation emphasizes the importance of individual differences among patients and suggests that theoretical and computational accounts of conceptual knowledge and its deterioration in semantic dementia would greatly benefit from future investigations looking at individual patients' brain scans in combination with the best possible information about premorbid biological and experiential factors, relating both region-specific atrophy and premorbid factors to performance in a range of tasks requiring knowledge of words and objects.

Manuscript received 21 May 2007
Revised manuscript received 12 September 2007
Revised manuscript accepted 2 October 2007
First published online 10 January 2008

REFERENCES

Blazely, A., Coltheart, M., & Casey, B. J. (2005). Semantic impairment with and without surface dyslexia: Implications for models of reading, *Cognitive Neuropsychology*, *22*, 695–717.

Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, *38*, 1207–1215.

Brunswick, N., McCrory, E., Price, C. J., Frith, C. D., & Frith, U. (1999). Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: A search for Wernicke's Wortschatz? *Brain*, *122*, 1901–1917.

Burnage, G. (1990). CELEX English lexical user guide. In G. Burnage (Ed.), *CELEX—a guide for users*. Nijmegen, The Netherlands: University of Nijmegen, Centre for Lexical Information.

Chan, D., Fox, N. C., Scahill, R. I., Crum, W. R., Whitwell, J. L., Leschziner, G., et al. (2001). Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. *Annals of Neurology*, 49, 433–442.

Cipolotti, L., & Warrington, E. K. (1995). Semantic memory and reading abilities: A case report. *Journal of the International Neuropsychological Society*, *1*, 104–110.

Coccia, M., Bartolini, M., Luzzi, S., Provinciali, L., & Lambon Ralph, M. A. (2004). Semantic memory is an amodal, dynamic system: Evidence from the interaction of naming and object use in semantic dementia. *Cognitive Neuropsychology*, *21*, 513–527.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel-distributed processing account of the Stroop effect. *Psychological Review*, *97*, 332–361.

Coltheart, M. (2004). Are there lexicons? *Quarterly Journal of Experimental Psychology*, *57*, 1153–1171.

Cortese, M. J., Balota, D. A., Sergent-Marshall, S. D., Buckner, R. L., & Gold, B. T. (2006). Consistency and regularity in past-tense verb generation in healthy ageing, Alzheimer's disease, and semantic dementia. *Cognitive Neuropsychology*, *23*, 856–876.

Dell, G. S., Martin, N., & Schwartz, M. F. (2007). A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming. *Journal of Memory and Language*, *56*, 490–520.

Funnell, E. (1996). Response biases in oral reading: An account of the co-occurrence of surface dyslexia and semantic dementia. *Quarterly Journal of Experimental Psychology*, *49A*, 417–446.

Galton, C. J., Patterson, K., Graham, K., Lambon Ralph, M. A., Williams, G., Antoun, N., et al. (2001). Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. *Neurology*, *57*, 216–225.

Garrard, P., & Carroll, E. (2006). Lost in semantic space: A multi-modal, non-verbal assessment of feature knowledge in semantic dementia. *Brain*, *129*, 1152–1163.

Graham, K. S., Hodges, J. R., & Patterson, K. (1994). The relationship between comprehension and oral reading in progressive fluent aphasia, *Neuropsychologia*, *32*, 299–316.

Graham, N. L., Patterson, K., & Hodges, J. R. (2000). The impact of semantic memory impairment on spelling: Evidence from semantic dementia. *Neuropsychologia*, *38*, 143–163.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meaning of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*, 662–720.

Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: Implications for the organization of semantic memory. *Memory*, *3*, 463–495.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain*, *115*, 1783–1806.

Howard, D., & Patterson, K. (1992). *The Pyramids and Palm Trees Test: A test of semantic access from words and pictures*. Bury St. Edmunds, UK: Thames Valley Test Company.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lambon Ralph, M. A., & Howard, D. (2000). Gogi aphasia or semantic dementia? Simulating and assessing poor verbal comprehension in a case of progressive fluent aphasia. *Cognitive Neuropsychology*, *17*, 437–465.

Lambon Ralph, M. A., McClelland, J. L., Patterson, K., Galton, C. J., & Hodges, J. R. (2001). No right to speak? The relationship between object naming and semantic impairment: Neuropsychological evidence and a computational model. *Journal of Cognitive Neuroscience*, *13*, 341–356.

Luzzi, S., Snowden, J. S., Neary, D., Coccia, M., Provinciali, L., & Lambon Ralph, M. A. (2007). Distinct patterns of olfactory impairment in Alzheimer's disease, semantic dementia, frontotemporal dementia, and corticobasal degeneration. *Neuropsychologia*, *45*, 1823–1831.

McCrory, E. J., Mechelli, A., Frith, U., & Price, C. J. (2005). More than words: A common neural basis for reading and naming deficits in developmental dyslexia? *Brain*, *128*, 261–267.

McKay, A., Castles, A., Davis, C., & Savage, G. (2007). The impact of progressive semantic loss on reading aloud. *Cognitive Neuropsychology*, *24*, 162–186.

Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S. J., & Hodges, J. R. (2000). A voxel-based study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, *47*, 36–45.

Patterson, K., & Hodges, J. (1992). Deterioration of word meaning: Implications for reading. *Neuropsychologia*, *30*, 1025–1040.

Patterson, K., & Hodges, J. (2000). Semantic dementia: One window on the structure and organisation of semantic memory. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology: Memory and its disorders* (2nd ed., Vol. 2, pp. 313–333). Amsterdam: Elsevier Science.

Patterson, K., Lambon Ralph, M. A., Hodges, J. R., & McClelland, J. L. (2001). Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. *Neuropsychologia*, *39*, 709–724.

Patterson, K., Lambon Ralph, M. A., Jefferies, E., Woollams, A., Jones, R., Hodges, J. R., et al. (2006). "Pre-semantic" cognition in semantic dementia: Six deficits in search of an explanation. *Journal of Cognitive Neuroscience*, *18*, 169–183.

Patterson, K., Suzuki, T., Wydell, T., & Sasanuma, S. (1995). Progressive aphasia and surface alexia in Japanese. *Neurocase, 1,* 155–165.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, *12*, 767–808.

Plaut, D. C. (2002). Graded modality-specific specialization in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology*, *19*, 603–639.

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

Pugh, K. R., Mencl, W. E., Jenner, A. R., Lee, J. R., Katz, L., Frost, S. J., et al. (2001). Neuroimaging studies of reading development and reading disability. *Learning Disabilities Research & Practice*, *16*, 240–249.

Rogers, T. T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M. L., Patterson, K., et al. (2006). The anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective, and Behavioral Neuroscience*, *6*, 201–213.

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., et al. (2004a). The structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, *111*, 205–235.

Rogers, T. T., Lambon Ralph, M. A., Hodges, J. R., & Patterson, K. (2004b). Natural selection: The impact of semantic impairment on lexical and object decision. *Cognitive Neuropsychology*, *21*, 331–352.

Rooney, P., Hussar, W., Planty, M., Choy, S., Hampden-Thompson, G., Provasnik, S., et al. (2006). *The condition of education*. Washington, DC: National Center for Education Statistics.

Rosen, H. J., Gorno-Tempini, M. L., Goldman, W. P., Perry, R. J., Schuff, N., Weiner, M., et al. (2002). Patterns of brain atrophy in frontotemporal dementia and semantic dementia. *Neurology*, *58*, 198–208.

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*, 228–264.

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

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

Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia: A form of circumscribed cerebral atrophy. *Behavioural Neurobiology*, *2*, 167–182.

Snowden, J. S., Neary, D., & Mann, D. M. A. (1996). *Frontotemporal lobar degeneration: Frontotemporal dementia, progressive aphasia, semantic dementia.* New York: Churchill Livingstone.

Snowden, J. S., Thompson, J. C., & Neary, D. (2004). Knowledge of famous faces and names in semantic dementia. *Brain*, *127*, 860–872.

Studholme, C., Cardenas, V., Blumenfeld, R., Schuff, N., Rosen, H. J., Miller, B., et al. (2004). Deformation tensor morphometry of semantic dementia with quantitative validation. *NeuroImage*, *21*, 1387–1398.

Welbourne, S. R., & Lambon Ralph, M. A. (2005). Exploring the impact of plasticity-related recovery after brain damage in a connectionist model of single-word reading. *Cognitive, Affective, & Behavioral Neuroscience*, *5*, 77–92.

Whitwell, J. L., Anderson, V. M., Scahill, R. I., Rossor, M. N., & Fox, N. C. (2004). Longitudinal patterns of regional change on volumetric MRI in frontotemporal lobar degeneration. *Dementia & Geriatric Cognitive Disorders*, *17*, 307–310.

Woollams, A., Lambon Ralph, M. A., Plaut, D. C., & Patterson, K. (2007). SD-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review*, *114*, 316–339.

APPENDIX A

Visual patterns and action patterns generation

Algorigthm

A total of 240 binary patterns were generated by giving a 1 with probability of .8 at positions marked with " + ", probability .2 at positions marked with "0", and probability 0 at positions marked with "−".

Visual Prototypes

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
living	animal	mammal	+	+	+	+	0	0	0	0	0	+	+	+	0	+	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
		reptile	+	+	+	+	0	0	0	0	0	+	+	+	+	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
		bird	+	+	+	+	0	0	0	+	+	-	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	plants	fruits	+	0	0	0	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+	+	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0
		vegetables	+	0	0	0	+	+	+	-	-	-	-	-	-	-	-	-	-	-	0	0	+	+	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0
non-living	musical instruments		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	0	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	+	+	+	+	0	0	0	0
	clothes		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	0	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	+	+	+	+	0	0	-	-	0	0	0	0
	utensils		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	0	0	0	-	-	0	0	-	-	0	0	-	-	+	0	0	-	+	+	+	+	0	0	-	-	0	0	-	-	0	0	0	0
	tools		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	0	0	0	-	-	0	0	-	-	0	0	-	-	+	+	+	+	+	0	0	-	-	0	0	-	-	0	0	-	-	0	0	0	0
	appliances		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	0	0	0	-	-	0	0	-	-	+	+	+	+	0	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	0	0
	transport		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	0	0	-	-	+	+	+	+	0	0	-	-	0	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	0	0
	furniture		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	0	0	-	-	0	0	-	-	0	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	0	0

Action Prototypes

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
living	animal	mammal	+	+	0	-	-	+	0	+	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
		reptile	+	+	0	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
		bird	+	+	0	-	-	+	0	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	plants	fruits	+	-	-	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
		vegetables	+	-	-	+	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
non-living	musical instruments		-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	clothes		-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	+	+	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	utensils		-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	tools		-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	+	+	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	appliances		-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	0	0	0	-	-	-	-	-	-	-	-	0	0
	transport		-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	0	0	-	-	-	-	0	0
	furniture		-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	0	0	0	0

APPENDIX B

Phonological patterns and orthographic patterns generation

Algorithm

The CVCC patterns were generated by giving 1 to the vowel marked with "+" and 0s to all other vowels (marked with "−"). Onsets and codas were picked for each item by giving each of the 12 consonants (marked with "0") equal chance (i.e., probability of being selected = 1/12) at each of the three positions (one for the onset, two for the coda). Once the consonants were picked, they were given a 1 while all others were given 0s. Thus, all 240 patterns consisted of 36-item long vectors with only four 1s (one for onset in the first 12 positions, one for vowel in the second 12 positions, and two for coda in the last 12 positions). The 240 unique orthographic patterns were first generated and then matched with the appropriate phonological pattern. The numbers in the first column refer to the number of patterns with that specific vowel.

Phonological Prototypes

no of	onset												vowel												coda											
patterns	b	p	d	t	g	k	m	n	l	f	s	ś	ej	æ		aj	ɪ		ow	o		ij	ɛ		b	p	d	t	g	k	m	n	l	f	s	ś
37	0	0	0	0	0	0	0	0	0	0	0	0	-	+	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	+	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	+	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	+	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	+	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	+	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	+	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	+	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	+	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	+	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	+	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	+	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	+	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	+	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	+	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	+	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	+	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	+	-	-	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	+	-	-	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	+	-	-	0	0	0	0	0	0	0	0	0	0	0	0

Orthographic Prototypes

no of	onset												vowel												coda											
patterns	b	p	d	t	g	k	m	n	l	f	s	sh	a	A	ai	i	l	Y	o	O	oa	e	ea	ie	b	p	d	t	g	k	m	n	l	f	s	sh
37	0	0	0	0	0	0	0	0	0	0	0	0	+	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	+	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	+	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	-	+	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	-	-	+	-	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	+	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	+	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	+	-	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	+	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	+	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	+	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	+	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	+	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	+	-	-	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	+	-	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	+	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	+	-	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	+	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	+	-	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	+	0	0	0	0	0	0	0	0	0	0	0	0

APPENDIX C

Full list of successful fits for the five patients

		Successful fit parameters			*Statistics*	
Patient	*No. fits*	*Training O:V*	*Direct pathway size*[a]	*Lesion bias*	χ^2	*p*
J.L.	43	1:1	10	unbiased	3.567	.468
				50% vis	0.309	.989
				75% vis	2.300	.681
				100% vis	5.302	.258
			20	100% orth	2.591	.628
				75% orth	2.264	.687
				50% orth	0.120	.998
				unbiased	0.531	.970
				50% vis	4.333	.363
				75% vis	5.171	.270
			30	100% orth	2.591	.628
				75% orth	2.444	.655
				50% orth	1.348	.853
				unbiased	2.229	.694
				50% vis	5.171	.270
		1:2	10	unbiased	6.171	.187
				50% vis	3.818	.431
				75% vis	4.393	.355
				100% vis	5.099	.277
			20	75% orth	5.636	.228
				50% orth	6.171	.187
				unbiased	4.660	.324
				50% vis	0.103	.999
				75% vis	3.256	.516
				100% vis	4.094	.394
			30	75% orth	3.074	.546
				50% orth	2.261	.688
				unbiased	0.872	.929
				50% vis	3.152	.533
				75% vis	3.185	.527
		2:1	10	100% orth	5.737	.220
				75% orth	4.127	.389
				50% orth	2.351	.672
				unbiased	2.052	.726
				50% vis	0.924	.921
			20	100% orth	1.482	.830
				75% orth	0.662	.956
				50% orth	1.501	.826
				unbiased	7.749	.101
				50% vis	7.749	.101
			30	100% orth	0.545	.969
				75% orth	3.066	.547
				50% orth	6.601	.159

(*Continued overleaf*)

Appendix C (Continued)

		Successful fit parameters			*Statistics*	
Patient	*No. fits*	*Training O:V*	*Direct pathway size*[a]	*Lesion bias*	χ^2	*p*
P.C.	22	1:1	10	50% vis	5.453	.244
				75% vis	5.494	.240
				100% vis	5.784	.216
			20	50% orth	4.481	.345
				unbiased	5.494	.240
				50% vis	5.856	.210
				75% vis	4.929	.295
			30	50% orth	5.065	.281
				unbiased	4.481	.345
		1:2	10	50% vis	6.704	.152
				75% vis	6.797	.147
				100% vis	4.481	.345
			20	50% vis	3.724	.445
				75% vis	6.367	.173
		2:1	10	75% vis	8.457	.076
			20	100% orth	3.782	.436
				75% orth	6.077	.193
				50% orth	5.494	.240
				unbiased	4.929	.295
			30	100% orth	5.694	.223
				75% orth	4.481	.345
				50% orth	6.367	.173
G.C.	18	1:1	10	50% orth	9.155	.057
				unbiased	7.594	.108
				50% vis	4.472	.346
			20	100% orth	7.348	.119
				75% orth	5.930	.204
				50% orth	7.186	.126
		1:2	10	unbiased	6.746	.150
			20	unbiased	8.321	.081
				50% vis	5.477	.242
			30	unbiased	8.985	.062
		2:1	10	75% orth	9.487	.050
				unbiased	9.150	.057
				50% vis	8.335	.080
			20	100% orth	8.671	.070
				75% orth	5.477	.242
				50% orth	7.702	.103
			30	100% orth	6.133	.189
				75% orth	7.282	.122
E.M.	15	1:1	30	50% vis	6.461	.167
				75% vis	5.379	.251
				100% vis	3.091	.543
		1:2	30	75% vis	7.478	.113
				100% vis	7.478	.113

(*Continued overleaf*)

Appendix C (Continued)

		Successful fit parameters			*Statistics*	
Patient	*No. fits*	*Training O:V*	*Direct pathway size*[a]	*Lesion bias*	χ^2	*p*
		2:1	10	75% vis	8.555	.073
				100% vis	7.469	.113
			20	unbiased	4.173	.383
				50% vis	6.461	.167
				75% vis	2.141	.710
				100% vis	2.141	.710
			30	unbiased	6.461	.167
				50% vis	4.178	.383
				75% vis	4.178	.383
				100% vis	3.091	.543
F.M.	7	1:1	30	75% vis	6.072	.194
				100% vis	4.865	.302
		2:1	20	75% vis	4.327	.364
				100% vis	4.485	.344
			30	50% vis	4.865	.302
				75% vis	3.408	.492
				100% vis	2.646	.619

Note: O:V = orthographic-to-visual ratio. vis = visual. orth = orthographic.