LETTER Communicated by Gary Cottrell

The Computational Exploration of Visual Word Recognition in a Split Model

Richard Shillcock
Department of Psychology and Institute for Adaptive and Neural Computation, Division of Informatics, University of Edinburgh, Scotland, U.K.

Padraic Monaghan
Institute for Adaptive and Neural Computation, Division of Informatics, University of Edinburgh, Scotland, U.K.

We argue that the projection of the visual field to the cortex constrains and informs the modeling of visual word recognition. On the basis of anatomical and psychological evidence, we claim that the higher-level cognition involved in word recognition does not completely transcend initial foveal splitting. We present a schematic connectionist model of word recognition that instantiates the precise splitting of the visual field and the contralateral projection of the two hemifields. We explore the special nature of the exterior (i.e., first and last) letters of words in reading. The model produces the correct behavior spontaneously and robustly. We analyze this behavior of the model with respect to words and random patterns and conclude that the systematic division of the visual input has predictable, general informational consequences and is chiefly responsible for the exterior letters effect.

1 Introduction

The human fovea is precisely split about a vertical midline: the left and right hemifields are projected contralaterally to the right and left hemispheres of the brain, respectively (Fendrich & Gazzaniga, 1989; Fendrich, Wessinger, & Gazzaniga, 1996; Sperry, 1968). This foveal splitting is sufficiently exact to mean that when a printed word is fixated, under typical reading distances, the two parts of the word are initially projected to different hemispheres (Sugishita, Hamilton, Sakuma, & Hemmi, 1994). Foveal splitting has been recognized in cognitive neuropsychological research, particularly research involving split-brain subjects, but it has received little attention in psycholinguistic approaches to visual word recognition (Brysbaert, 1994; Shillcock & Monaghan, 1998; Shillcock, Ellison, & Monaghan, in press; Shillcock, Monaghan, & Ellison, 1999). We describe a series of studies using neural network architectures to explore an abstract characterization of one aspect of visual word recognition: that of coordinating the information in the two

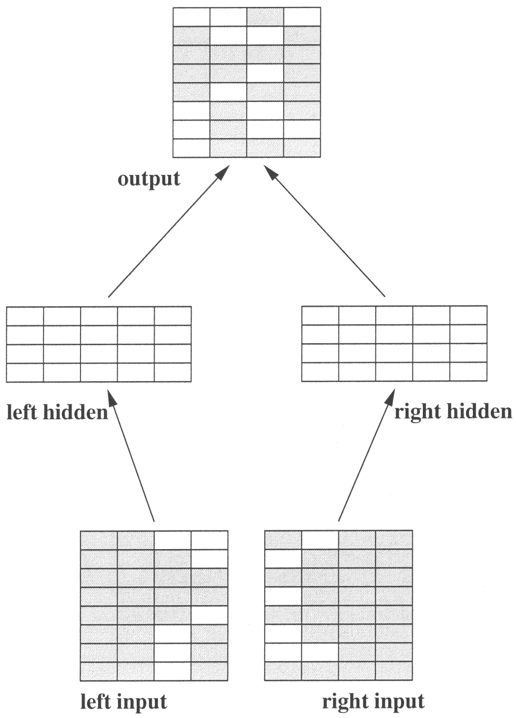

Figure 1: Split model architecture. Each column of eight units in the input and output layers represents one letter. Gray units denote nonactivated nodes; white units denote activated nodes. Words were presented in every position across the input units.

hemifields, containing, respectively, the two parts of a fixated word. We will show that one particular aspect of visual word recognition, the special status accorded to the exterior (i.e., first and last) letters of a word, may be explained by the split nature of the processor. More generally, we will show that such modeling can yield theoretical insights into the nature of hemispheric interaction.

Figure 1 illustrates a neural network model of word recognition that instantiates foveal splitting in the simplest, most extreme form. The model carries out an identity mapping from an orthographic representation presented across the input nodes to the same representation across the output nodes. The division between the two halves of the input is precise, with no sharing of information between the two hemifields. In reality, in the human visual system, it is likely that some degree of ipsilaterally projected information is available. For instance, low spatial-frequency information about

w	o	r	d
	w	o	r
		w	o
			w

d			
r	d		
o	r	d	
w	o	r	d

Figure 2: Example of the five different versions of the stimulus for any one word.

a fixated word may be accessible from subcortical routes, which are not dependent on callosal transfer; thus, information about the length of the fixated word may be available to both hemispheres. Elsewhere we argue in more detail that this initial splitting of the input conditions the higher cognitive processing involved in visual word recognition (Shillcock et al., in press; Shillcock & Monaghan, 1998; Shillcock et al., 1999). Our strategy has been first to investigate the robust behavior of the simplest model of foveal splitting. The critical feature of the model is that it is required to coordinate the information in the two hemifields regardless of where the fixation point falls. Thus, as shown in Figure 2, for a four-letter word there are five possible fixation points, ranging from just before the first letter to just after the last letter and ignoring the possibility of a fixation directly on a letter.

2 The First and Last Letters of Words

In psycholinguistic experiments, the first and last letters of a visually presented word have been shown to be more salient than the rest of the letters and to receive priority in processing. Forster (1976) argues that these exterior letters constitute an access code that activates a subset of the lexicon.[1] Rumelhart and McClelland (1981) suggest that "subjects use some sort of 'outside in' processing strategy that leads to variations in performance across serial position" (p. 76). Jordan (1990) concludes that "psychological representations for exterior letter combinations from words do appear to exist, and . . . can be activated even though no other letters are perceived" (p. 903). The priority of exterior letters is evident from a number of experimental paradigms:

- *Identity priming.* Recognition of the whole target word (e.g., trap) is facilitated by the prior presentation of its exterior letters (*t p*) but not its interior letters (*ra*) (Forster & Gartlan, 1975; McCusker, Gough, & Bias, 1981).

[1] Throughout this article we use Jordan's (1990) terminology of *exterior letters* to refer to the first and last letter of a visually presented word and *interior letters* to refer to all the letters that lie between the first and last position.

- *Probed report*. Exterior letters of pattern postmasked strings are reported more accurately than interior letters (Butler & Merikle, 1973; Merikle, 1974; Merikle & Coltheart, 1972; Merikle, Coltheart, & Lowe, 1971). Importantly, the two exterior letters produce closely comparable levels of report. Rumelhart and McClelland (1981) present data from the Reicher-Wheeler task, showing serial position effects in the form of a bow-shaped curve (experiment 7, figure 18, p. 77); it is these data that prompt Rumelhart and McClelland to make the "outside in" processing claim quoted above.

- *Identification*. Legal pairs of exterior letters such as (*d k*) (from *dark*, *disk*) are reported more accurately than single letters (*d*) or illegal pairs of exterior letters (*d x*) or (*z k*) (Jordan, 1990, 1995). This "pair-letter effect" was observed only when the pattern postmask matched the horizontal boundaries of the legal exterior letter pair. Reports of whole-word displays were always superior to reports of letter pairs.

- *Off-line identification*. Exterior letters tend to be recognized first in noisy conditions, as when successive presentations of blurred content words become clearer and clearer (see, e.g., Shillcock, Kelly, Buntin, & Patterson, 1997).

In summary, the exterior letters of visually presented words are afforded priority in processing in a variety of recognition tasks. One potential explanation is that although the exterior letters are in fact farthest away from the central foveal fixation point, the relevant visual information is actually clearer than that in the middle because exterior letters are bounded on one side by white space and are thus not susceptible to the same level of visual ambiguity and lateral interference as interior letters (Eriksen & Rohrbaugh, 1970; Estes, Allmeyer, & Reder, 1976). This explanation is perhaps most compelling for the off-line recognition of degraded lexical stimuli. Jordan (1990) argues against this lateral interference account as the sole explanation of all the data, on the grounds that not all types of string elicit the effect: pattern-postmasked strings of letter-like nonsense characters elicit the reverse pattern of results, with exterior characters being perceived less well than interior characters (Hammond & Green, 1982; Mason, 1982; Mason & Katz, 1976). These particular data are seen as resulting from a word recognition device specifically attuned to exterior letters, which is partially set in motion by the nonsense characters and cannot be completely inhibited. These data involving letter-like nonsense characters are simultaneously evidence for the special status of exterior letters and the claim that this status does not derive solely from more peripheral, psychophysical considerations.

The credit for the original demonstration that the effect can emerge from an implemented computational model goes to Rumelhart and McClelland, who suggest a number of mechanisms that might account for the serial position effects they observe in the human data; they implement two of them to simulate the data. First, they differentially weight the inputs to each of the four letter positions in their interactive-activation model (IAM). Second, they assume that the readout occurs at different times for the different positions: the two exterior letters are read out first, followed by the second and then the third. Rumelhart and McClelland also mention possible perceptibility differences of particular letters in particular positions, statistical properties of the words, and variations in locus of fixation and attention, thus summarizing the most relevant parameters in a monolithic (nonsplit) model of word recognition. The statistical properties of the English lexicon will be relevant to our own explorations below. The beginnings of English words are typically the most informative parts; redundancy rises across serial position (see, for instance, Yannakoudakis & Hutton, 1992, for an analysis of the phonological statistics of English words). In monosyllabic English words, the dominant consonant-vowel-consonant (CVC) structure allows greater variety in onsets and codas than in the middles of words. Thus, giving processing priority to the exterior letters of English words is adaptive.

The exterior letters effect (ELE) is outside the range of data that has been captured by models of visual word recognition in the connectionist tradition other than the IAM. In Seidenberg and McClelland's (1989) developmental model, the Wickelgraph input representation leads to the reverse prediction by underrepresenting the exterior letters compared with the interior letters, in its triples of consecutive letters: $^{*}da$, *dar*, *ark*, rk^{*} (this criticism also applies to Mozer's, 1987, BLIRNET model).[2] In some of the subsequent models of visual word processing (chiefly concerned with pronunciation) in this tradition (e.g., Plaut & McClelland, 1993), the orthographic input is explicitly structured in terms of onset, nucleus, and coda, and there is no indication that any differential processing can emerge for exterior letters; indeed, Plaut and McClelland remark on the typically autonomous and componential processing occurring in onset, nucleus, and coda, the three positions interacting in the processing of only words with irregular pronunciations, like *pint*. In summary, existing models of visual word recognition do not spontaneously allow the ELE to emerge as a principled result of their basic architecture. Below, we develop an account of the effect, based on the claim that foveal splitting fundamentally conditions word recognition. We present simulations using a lexicon of the 60 most frequent English words to demonstrate that the ELE occurs with psychologically realistic stimuli, and we present simulations using more controlled lexica to establish that the effect is genuinely due to the split nature of the processor.

[2] The computational measure of coding ends of words in two ways, as * and **, allows interior and exterior letters to participate in an equal number of representations: $^{**}d$, $^{*}da$, *dar*, *ark*, rk^{*}, k^{**}. However, this amendment still does not make the correct prediction of better representation of exterior letters.

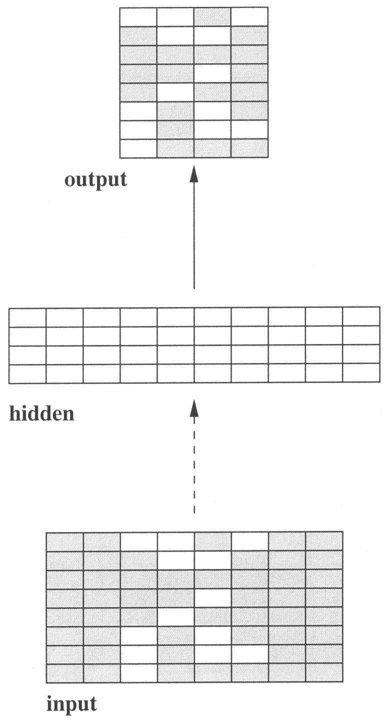

Figure 3: Nonsplit model. The dashed line indicates incomplete connectivity between the input and the hidden layers. See the text for details.

We compare the behavior of the split model with an otherwise comparable nonsplit model, shown in Figure 3. In the simulations with a nonsplit architecture, we used the same training and testing regime as with the split architecture. In the nonsplit model, the input layer consisted of eight letter positions, effectively pooling the separate input layers of the split model. The hidden layer was also a pooled-resource version of the split model, and contained 40 units. The hidden layer and the output layer were fully connected. However, in order to make the processing power of the two networks comparable, each unit in the input layer in the nonsplit model was connected randomly to only half the hidden units. This measure allows there to be the same number of weighted connections between input and hidden layers in the split and nonsplit models. The processing power of neural networks is a function of the number of units and the number of weighted connections. If the two models are alike in these respects, then their performances may be compared.

3 Simulating the Exterior Letters Effect

Figure 1 shows the basic model that we explored. The input layer is composed of two sets of four input units (four letter slots) on either side of a midline, or fixation point. There is complete connectivity between these sets of input units and their respective groups of 20 hidden units. The number of hidden units was determined by pilot studies and represents the minimum number capable of reliably solving this mapping problem. There is complete connectivity between the two sets of hidden units and the set of four output units. Our network encodes letters in terms of the eight-bit visual feature system employed by Plaut and Shallice (1994). The features correspond to attributes such as “contains a vertical stroke” or “contains a closed part.” Each input layer and the output layer contain four letter positions. The model was required to recreate and integrate its orthographic inputs at the output units, for all five possible inputs for each word (see Figure 2). The network was trained using the backpropagation learning algorithm (Rumelhart, Hinton, & Williams, 1986) and was implemented in PDP++ (O’Reilly, Dawson, & McClelland, 1995).

The principal psychological reality that this model is intended to capture is that of the splitting of the visual field. The set of output units is not necessarily intended to be located in any one hemisphere. We use the model to address the question of what happens when information divided between two processing domains is required to be coordinated. What processing strategies spontaneously emerge from a simple instantiation of the problem in a split architecture?

3.1 Simulating the ELE with a Real-Word Lexicon. The exterior letters of four-letter words, such as *d**k* as in *disk*, are recognized better than the interior letters of words in studies with human subjects (Forster & Gartlan, 1975; McCusker et al., 1981). In this simulation we aim to show that this effect naturally emerges from a split model trained on a small but realistic set of English words. The 60 words used in the training set, and listed in the appendix, were all the four-letter words with a frequency greater than 1 in 10,000 from the CELEX lexical database (Baayen, Pipenbrock, & Gulikers, 1995). Each word was presented to the model an equal number of times in each of the five possible input locations. For each presentation, the model was required to recreate the word at the output. The words were presented in random permuted order, with particular presentation positions also randomized. After approximately 100 epochs of training, the network had learned the task well, with mean squared error (MSE) for each word being less than 0.5. The nonsplit network was similarly trained on the same lexicon, with the words presented in all five possible positions in the input and was trained to the same level of MSE as the split model. Ten simulations were carried out, and all produced very similar results.

We tested the hypothesis that the ELE would emerge within these simple constraints. The trained models were presented with stimuli consisting of either the exterior or the interior letters of each of the words in the training set. These letters were positioned so that the notional words of which they were a part occupied all possible positions across the input units (below, we refer to the individual input nodes by number, counting from left to right from 1 to 8). In the exterior letters condition, the exterior letters of each word in the training set were presented intact, and each of the interior letters was replaced by an ambiguous pattern in which all the units were activated to half their maximum value.[3] We used these ambiguous patterns to represent stimulus conditions in the original experiments by Jordan and others in which information was provided about word length but with no indication about what letters were present in those particular positions.[4] In the interior letters condition, the interior letters were presented intact, and the exterior letters were replaced by ambiguous patterns.

Figure 4 shows the MSE at output letter positions 3 and 6 (the exterior letters, given a central fixation of the word) for the split and nonsplit networks respectively, for the exterior letters condition. Figure 5 shows the MSE for the output letter positions 4 and 5 (the interior letters, given a central fixation of the word) for the split and nonsplit networks, respectively, for the interior letters condition. Taken together the two graphs show a striking interaction between model type (split versus nonsplit) and letters presentation condition (exterior versus interior): the difference between the MSE for exterior letters and interior letters is greater in the split model than in the nonsplit model. The graphs also show a somewhat smaller MSE for the exterior letters compared with the interior letters in the nonsplit model alone. An analysis of variance was carried out treating the individual trained models as subjects in an experiment; the MSE was summed for the two interior letters and the two exterior letters in each model. The analysis of variance confirmed a significant interaction between model type and letters presentation condition ($F(1,18) = 31.23$, $p < .001$; $F(1,59) = 149.47$, $p < .001$). Individual *t*-tests were performed on the MSE summed across the two exterior letters and across the two interior letters for each model. For the split model, there was a significant difference between MSE for the exterior letters and MSE for the interior letters ($t(9) = -9.89$, $p < .001$, two-tailed). For the exterior letters alone, there was a significant difference between the split model and the nonsplit model ($t(9) = -6.35$, $p < .001$, two-tailed). For the interior letters alone, there was no significant difference between the split model and the nonsplit model ($t(9) = .39$, n.s.). These results all represent an ELE in which the split nature of one of the models plays a crucial part. However, there is

[3] This pattern of activation is exactly equidistant from each letter in the eight-dimensional input space.

[4] In the experimental paradigm used by Jordan (1990, 1995), information about word length was conveyed by the stimulus mask.

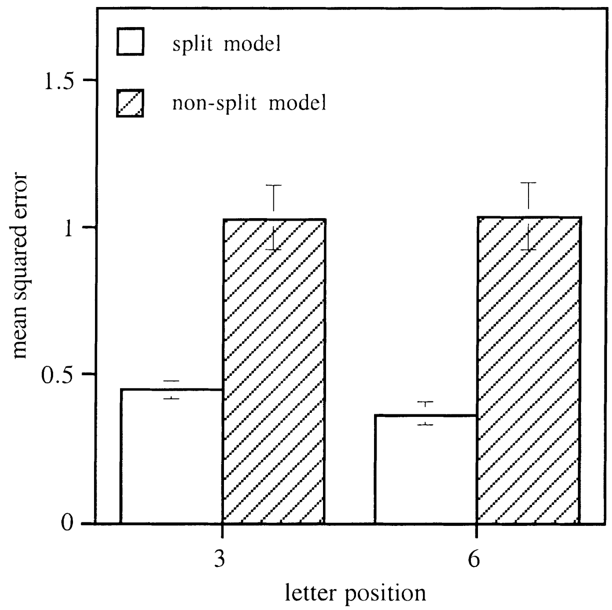

Figure 4: Mean squared error (MSE) at individual letter positions 3 and 6 for the 60 words of English when the exterior letters of the words were presented as input to the split and nonsplit model, respectively. Error bars in all graphs represent standard error of the mean.

also a (smaller) significant difference between the exterior and interior letters in the nonsplit model alone ($t(9) = -2.80$, $p = .021$ two-tailed). This last result suggests an ELE with its origin in the basic task—the shift-invariant identity mapping—that the models are both required to perform.

These results refer only to the central fixation, across input nodes 3, 4, 5, and 6. Separate analyses were carried out for each of the five possible input positions for a word. The results are presented in Tables 1, 2, and 3; the middle rows of the tables contain the results discussed above, and illustrated in Figures 4 and 5, for the central fixation.

Table 1 shows the MSE data for the exterior letters condition and Table 2 for the interior letters condition. Table 3 summarizes the results of analyses of variance summing MSE for the two interior and the two exterior letters in each model, with subjects (F(1,18)) and items (F(1,59)), respectively, as the random variable; only the outcome of the (model $\times$ presentation condition) interaction is shown, in the second column. The four right-most columns of Table 3 show the results of the *t*-tests (df = 9) between summed MSE for exterior and for interior letters and split and nonsplit models. The pattern of results described above for the central fixation is broadly followed in

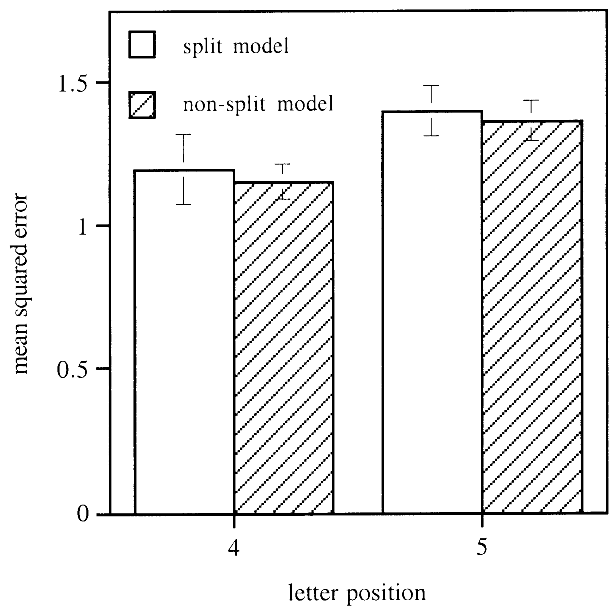

Figure 5: MSE at individual letter positions 4 and 5 for the 60 words of English when the interior letters of the words were presented as input to the split and nonsplit models, respectively.

Table 1: MSE for Presentation of the Exterior Letters to the Split and Nonsplit Models Trained on the 60 Most Frequent English Four-Letter Words.

Model	Mean Squared Error in Letter Position							
	1	2	3	4	5	6	7	8
Split	0.68	-	-	0.93				
Nonsplit	0.85	-	-	0.89				
Split		0.50	-	-	0.19			
Nonsplit		1.04	-	-	0.95			
Split			0.45	-	-	0.37		
Nonsplit			1.03	-	-	1.04		
Split				0.20	-	-	0.49	
Nonsplit				1.07	-	-	1.01	
Split					0.97	-	-	0.65
Nonspit					0.85	-	-	0.78

the two other more central presentations, in which the word is "fixated" between two of its constituent letters and input falls into each half of the model; the interaction was significant in each. In single hemifield presen-

Table 2: MSE for Presentation of the Interior Letters to the Split and Nonsplit Models Trained on the 60 Most Frequent English Four-Letter Words.

Model	Mean Squared Error in Letter Position							
	1	2	3	4	5	6	7	8
Split	-	0.81	1.21	-				
Nonsplit	-	0.80	1.13	-				
Split		-	1.23	1.64	-			
Nonsplit		-	0.92	1.27	-			
Split			-	1.19	1.40	-		
Nonsplit			-	1.15	1.36	-		
Split				-	1.19	1.42	-	
Nonsplit				-	1.13	1.37	-	
Split					-	0.93	1.36	-
Nonspit					-	0.92	1.22	-

Table 3: Results of the Analyses of Variance and *t*-Tests (Two-Tailed) for the Split and Nonsplit Networks Presented with the 60 Most Frequent Words of English.

Input Nodes	Interaction Between Model and Presentation Condition, by Subjects and by Items	*t*-Test for Exterior Letters, Comparing Split and Nonsplit Models	*t*-Test for Split Model, Comparing Exterior and Interior Letters	*t*-Test for Interior Letters, Comparing Split and Nonsplit Models	*t*-Test for Nonsplit Model, Comparing Exterior and Interior Letters
1234	n.s.,***	n.s.	*	n.s.	n.s.
2345	***,***	***	***	*	n.s.
3456	***,***	***	***	n.s.	*
4567	***,***	***	***	n.s.	*
5678	n.s.,***	n.s.	***	n.s.	**

Note: $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

tations, in which all of the input falls in half of the model, the interaction becomes nonsignificant in the analyses by subjects. In some of the individual models trained from different random starting weights, the interactions for the single hemifield presentations achieved significance in the analyses by subjects, but the interaction was always weaker than those for the split (more central) presentations. Table 3 shows the precise role of the exterior letters in the split model in causing the interaction. The ELE seems to receive contributions from at least two sources: the split architecture of one of the models and the nature of the task itself. Each model is required to perform the same task, so the significant interaction between model type and presentation condition must reflect the architectural difference between the models—the fact that one of them is split. The data in the right-most

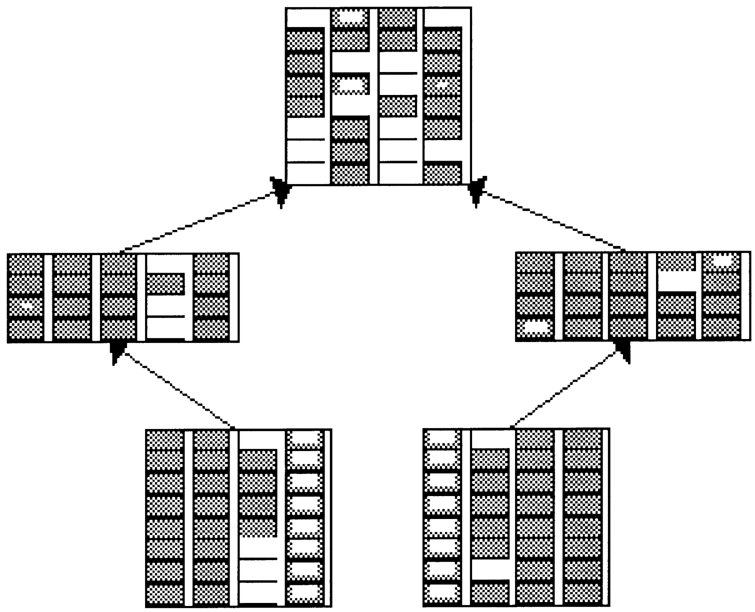

Figure 6: PDP++'s representation of the split network's performance on the centrally presented word *mean* when only the exterior letters are presented. Each node's activation level is indicated by the size of the light area; when the node is highly activated, it is wholly white, and when it is inactive, it is wholly gray.

column of Table 3 are evidence for a smaller potential contribution to the ELE from the task itself.

Figure 6 shows the PDP++ representation of the network when it is presented with only the exterior letters: *m****n* of the centrally presented word *mean*. The exterior letters are reproduced with very little error in the output. Figure 7 shows the network when it is presented with only the corresponding interior letters: **ea**. In this condition, the features of the interior letters are not reproduced accurately. For the letter *e*, one of the three features is not activated very strongly, and two features are erroneously activated. For the letter *a*, three of the four features are not activated strongly, and two features are activated erroneously. It is also informative to compare the activation of letters that are not presented in the input in Figures 6 and 7. When only the exterior letters are presented, the network correctly produces five of the seven features present in the interior letters and incorrectly activates four features. When only the interior letters are presented the network correctly produces four of the seven features of the exterior letters but inappropriately activates seven other features. This detailed study of one stimulus

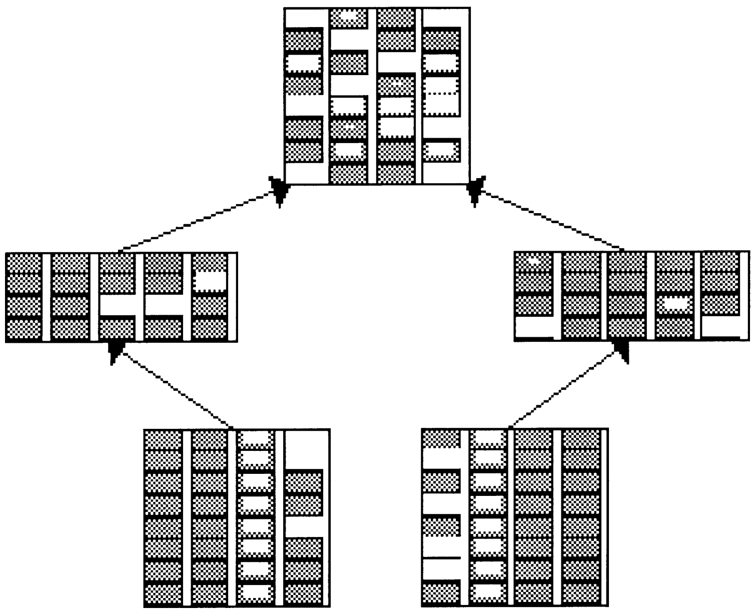

Figure 7: The network's performance on the word *mean* when only the interior letters are presented. Performance is worse than when only exterior letters are presented. See the text for details.

word, *mean*, presented in the two conditions provides a qualitative picture of the ELE as it appears in the whole-word priming task, which we describe in more detail below.

McCusker et al. (1981) show, in a series of four experiments, that the exterior letters of four-letter words are better than the corresponding interior letters at priming the subsequently presented whole word. We simulated this effect using the 10 different versions of each of the split and nonsplit model trained on the 60-word lexicon, as described above. The exterior letters or the interior letters of each word were presented, and the total MSE associated with the relevant whole word at the output was recorded. This measure indicates how close the network comes to reproducing the whole word from just the presented letters. We tested the hypothesis that the exterior letters would lead to a significantly smaller MSE for the whole word, compared with the MSE generated by the interior letters.

Two-way ANOVAs with model type (Split versus Nonsplit) and letters presentation condition (Exterior versus Interior) and MSE over the whole word as dependent variable were performed for each of the 10 split and nonsplit simulations. The effects were qualitatively identical to when error only for presented letters was analyzed, as above: presenting exterior letters to the split model primes the whole word better than presenting interior letters, whereas for the nonsplit model, this difference in priming is significantly smaller. For the centrally presented words (input nodes 3, 4, 5, and 6), for example, MSE with the split model was 5.70 when the exterior letters were presented and 8.22 when the interior letters were presented. For the nonsplit model, MSE was 7.27 following presentation of the exterior letters and 7.60 following presentation of the interior letters. For the single-hemifield presentations (input nodes 1, 2, 3, 4 and 5, 6, 7, 8) the interactions between model type (Split versus Nonsplit) and presentation condition (Exterior versus Interior) were not significant ($F(1, 18) = 0.16$, $p = 0.69$; and $F(1, 18) = 1.81$, $p = 0.20$, respectively). However, for the other presentation positions (input nodes 2, 3, 4, 5; 3, 4, 5, 6; and 4, 5, 6, 7), there were significant interactions ($F(1, 18) = 47.91$, $p < 0.001$; $F(1, 18) = 45.08$, $p < 0.001$; $F(1, 18) = 78.43$, $p < 0.001$, respectively). For the presentations that cross the midline of the split network, the ELE obtains in the simulation of the whole-word priming task.

We report this simulation as a direct demonstration of the modeling of psychological data related to the ELE. We hypothesize that the superior whole-word priming observed for the exterior letters is due to the split architecture of the processor, but support for this hypothesis must await the control studies with completely artificial lexica, which we report below. However, the simulations reported above show that the ELE is a robust effect with a real-word lexicon.

In summary, construing part of visual word recognition as a shift-invariant identity mapping and training with the 60 most frequent four letter words of English produce the ELE, and produce it significantly more robustly in the split model. Presenting just the exterior letters leads to good representations of those letters and hence to effective primes for words, whereas presenting just the interior letters produces less secure representations and hence less effective priming of the appropriate word. One factor potentially contributing toward such behavior is the particular statistics of the training set. As an inspection of the appendix reveals, interior letters are more ambiguous in the words to which they refer: there are 9 ambiguous exterior letter-pairs (e.g., *h**e* matches both *have* and *here*), including two three-way ambiguities (e.g., *life*, *like*, *late*), whereas there are 15 ambiguous interior letter pairs, including 4 three-way ambiguities (e.g., *keep*, *seem*, *feed*) and one four-way ambiguity (*them*, *then*, *when*, *they*). Exterior letters cue a smaller number of words than do interior letters, and it is therefore useful to prioritize their representation compared with the interior letters. The statistics of the training set broadly reflect the fact that in English CVC words, more variety is possible in the consonants of the onset and coda than in the vowels of the nucleus. A second potentially contributing explanation is that the observed ELE emerges from the split nature of one of the processors. A third potential explanation is that it is due to the nature of the shift-invariant identity mapping that the task requires. A fourth potential account of the ELE in the split model involves the definition of letters in terms of their immediate contexts: the orthotactic constraints of English monosyllabic words dictate, for instance, that *f* may be followed by *t*, but that the reverse is not allowed, and that only certain vowels in certain orders may appear adjacently. We may expect the nonsplit model as well as the split model to encode letter identity at least in part in terms of immediate context, so that an unspecified letter might be partially restored. In the split model, this means of encoding letter identity might be expected to militate against interior letters compared with exterior letters in that the contexts of the former are more adversely affected by the split and the shift-invariant identity mapping.

In the simulations described below, we address the four potential explanations given above by using lexica composed of items in which the degree of structure-mediating features and words in the original real-words lexicon is decreased. First, below, we randomize the order of letters within the words, thereby removing orthotactic structure from the lexicon. Further, we remove the letter level completely in a lexicon in which each word is represented by a unique random array of features. We hypothesize that the split nature of the processor will produce the most robust contribution to the ELE.

3.2 Simulating the ELE in a Random-Letter Lexicon. The random-letter lexicon contained 60 items. Items were generated by randomly assigning a letter to each position within the item. Randomization was permuted so that each letter occurred at least twice in each letter position across the lexicon. This random-letter lexicon allows us to assess the extent to which the ELE observed with real words was due to the orthotactic constraints present in real words of English.

The split and nonsplit models were trained and tested, as with the real-word lexicon above, by presenting first the exterior letters and then the interior letters and recording the respective MSE. Tables 4 and 5 present the data for all presentation positions, as in the analysis of the real-words lexicon above. Table 6 summarizes the model $\times$ presentation condition interaction from analyses of variance conducted as in the real-words simulation above. The interaction between model type and presentation condition was again the most striking aspect of the data. The ELE emerged robustly, as with the real-words lexicon. However, in the nonsplit model, the MSE for the exterior letters was also noticeably smaller than that for the interior letters. As Tables 4, 5, and 6 show, with the random-letter words, the interaction between model type and presentation condition reaches significance in both of the single hemifield presentations. The ELE obtains even when the word is being presented to a single hemifield.

Table 6 shows that in the simulations with the random-letter words, the overall pattern of results found for the real English words obtains more robustly. In the current simulation the model $\times$ presentation condition interaction is significant in the analyses by subjects even in the single-hemifield

Table 4: MSE for Presentation of the Exterior Letters to the Split and Nonsplit Models Trained on the Random-Letter Words.

Model	Mean Squared Error in Letter Position							
	1	2	3	4	5	6	7	8
Split	.84	-	-	1.14				
Nonsplit	1.05	-	-	1.19				
Split		.69	-	-	.27			
Nonsplit		1.34	-	-	1.33			
Split			.61	-	-	0.49		
Nonsplit			1.51	-	-	1.39		
Split				.28	-	-	.62	
Nonsplit				1.54	-	-	1.46	
Split					1.18	-	-	.92
Nonspit					1.37	-	-	1.23

Table 5: MSE for Presentation of the Interior Letters to the Split and Nonsplit Models Trained on the Random-Letter Words.

Model	Mean Squared Error in Letter Position							
	1	2	3	4	5	6	7	8
Split	-	1.64	1.55	-				
Nonsplit	-	1.36	1.46	-				
Split		-	1.69	1.71	-			
Nonsplit		-	1.65	1.57	-			
Split			-	1.93	1.73	-		
Nonsplit			-	2.04	1.72	-		
Split				-	2.05	1.80	-	
Nonsplit				-	1.73	1.64	-	
Split					-	1.56	1.38	-
Nonspit					-	1.48	1.34	-

presentations, and four of five of the MSE differences between the exterior letters in the split and nonsplit model are highly significant, the *t*-values being greater than the corresponding values in Table 3 for the real-words simulation. Similarly, in the right-most column of Table 6, more of the *t*-tests of the difference between MSE for the interior and exterior letters in the nonsplit model are significant, compared with the same comparisons in Table 3.

In summary, the simulations with the random-letter lexicon rule out two of the four potential accounts we suggested for the ELE observed in our simulations with the real-words lexicon. The orthotactic constraints present in the 60-word English lexicon could not have been wholly responsible for the effect. Similarly, the split architecture's greater disruption of the letter context of the interior letters compared with the exterior letters could not have

Table 6: Results of the Analyses of Variance and *t*-Tests (Two-Tailed) for the Split and Nonsplit Networks Presented with the 60 Random-Letter Words.

Input Nodes	Interaction Between Model and Presentation Condition, by Subjects and by Items	*t*-Test for Exterior Letters, Comparing Split and Nonsplit Models	*t*-Test for Split Model, Comparing Exterior and Interior Letters	*t*-Test for Interior Letters, Comparing Split and Nonsplit Models	*t*-Test for Nonsplit Model, Comparing Exterior and Interior Letters
1234	**,***	n.s.	***	**	**
2345	***,***	***	***	n.s.	***
3456	***,***	***	***	n.s.	***
4567	***,***	***	***	**	*
5678	**,***	**	***	n.s.	n.s.

Note: *$p < .05$. **$p < .01$. ***$p < .001$.

played a major role. The ELE was more pronounced in the random-letter lexicon, despite the absence of reliable orthotactic contexts. The ELE in both the real-words simulations and the random-letters simulations is caused by the nature of the processing task; it emerges spontaneously and robustly from the shift-invariant identity mapping and is significantly magnified by the presence of a split in the processor.

We repeated the simulations corresponding to whole-word priming with the random-letters lexicon. We tested the prediction that compared with the interior letters, the exterior letters would produce a significantly smaller MSE for the whole word of which they were part.

Two-way ANOVAs with model type (Split versus Nonsplit) and letters-presentation condition (Exterior versus Interior) and MSE over the whole word as dependent variable were calculated as above, and the effects were qualitatively similar. For the centrally presented words (input nodes 3, 4, 5, 6), for example, MSE for the split model was 5.90 when the exterior letters were presented and 9.87 when the interior letters were presented. For the nonsplit model, MSE was 8.31 following presentation of the exterior letters and 9.10 following presentation of the interior letters. For the single-hemifield presentations (input nodes 1, 2, 3, 4 and 5, 6, 7, 8), the interaction between model type (Split versus Nonsplit) and presentation condition (Exterior versus Interior) was significant in the right visual field presentation ($F(1, 18) = 6.34$, $p = 0.021$) and was marginally significant in the left visual field presentation ($F(1, 18) = 3.36$, $p = 0.083$). For the other presentation positions (input nodes 2, 3, 4, 5; 3, 4, 5, 6; and 4, 5, 6, 7) there were significant interactions ($F(1, 18) = 443.66$, $p < 0.001$; $F(1, 18) = 94.53$, $p < 0.001$; $F(1, 18) = 195.97$, $p < 0.001$, respectively). With lexical entries containing no orthotactic structure, the ELE obtains more strongly than in the same simulations with the real-word lexicon.

The results show that exterior letters are better primes than interior letters for the words from which they are drawn, even when the words are random strings of letters. These results support our interpretation of the same effect for the real-word lexicon that the effect was due to the nature of the processor and the task, as opposed to being solely due to the orthotactic structure contained in the real words or the unequal disruption of the immediate contexts of exterior and interior letters. The simulations show that the ELE in whole-word priming is a sufficiently strong effect not to be disrupted by the presence of realistic orthotactic structure in the training data.

3.3 Simulations with Random Feature Patterns as "Words". In the two sets of simulations, we have used real English words and artificial words composed of randomly scrambled letters, respectively. In both cases, there is a letter level of representation between the lexical and featural levels in the input lexicon, and the hidden units may be encoding generalizations at this letter level. The real-word input allowed intermediate generalizations about letters; for instance, 23 letters occurred in the training set, the letter *y* may occur only in particular locations, the sequence *rd* may occur word-finally but not word-initially, and the sequence *aa* is not permitted. The random-letters lexicon contained a letter level of representation although no orthotactic structure, and generalizations were possible concerning the within-letter contexts in which particular features could occur. We carried out a final set of simulations with an input lexicon with no such letter level. We created a lexicon of 60 random distributions of features, which may be thought of as single-character pictograms: each word is unique, and it does not share any predictable structure with any other words above the level of individual features. Thus, in Figure 1, the 4×8 feature map representing *word* was effectively scrambled, although each "letter" so produced was composed on average of the same number of features used in the real-word lexicon, and each "letter" consisted of at least one activated feature. This randomized input allowed generalizations only at the feature level (i.e., whether a particular node is activated) and at the "word," or pictogram, level (the whole 4×8 pattern). This final input lexicon therefore represents a more difficult mapping problem than the two previous lexica. We tested the hypothesis that the ELE would emerge robustly for this input with random patterns as words.

Simulations with the two models were carried out as in the two previous sets of simulations. The models learned to the same criterion as for the lexical input after approximately 150 epochs of training. The results for the central presentation are presented in Tables 7, 8, and 9, showing a yet clearer manifestation of the behavior seen in the two sets of simulations above. There is a strong interaction of model type and presentation condition and an ELE in the nonsplit model. Compared with the corresponding results for the real-word lexicon and the random-letters lexicon, the pattern of results in Table 9 shows increased uniformity in the model $\times$ presentation condition

Table 7: MSE for Presentation of the Exterior Letters to the Split and Nonsplit Models Trained on the 60 "Pictogram" Words.

Model	Mean Squared Error in Letter Position							
	1	2	3	4	5	6	7	8
Split	1.13	-	-	1.44				
Nonsplit	1.13	-	-	1.52				
Split		.90	-	-	.56			
Nonsplit		1.29	-	-	1.52			
Split			.85	-	-	.98		
Nonsplit			1.36	-	-	1.43		
Split				.48	-	-	1.01	
Nonsplit				1.44	-	-	1.42	
Split					1.40	-	-	1.20
Nonspit					1.23	-	-	1.11

Table 8: MSE for Presentation of the Interior Letters to the Split and Nonsplit Models Trained on the 60 "Pictogram" Words.

Model	Mean Squared Error in Letter Position							
	1	2	3	4	5	6	7	8
Split	-	1.67	1.95	-				
Nonsplit	-	1.69	1.81	-				
Split		-	1.96	2.17	-			
Nonsplit		-	1.93	1.94	-			
Split			-	2.24	2.11	-		
Nonsplit			-	2.19	2.18	-		
Split				-	2.02	2.02	-	
Nonsplit				-	2.04	2.01	-	
Split					-	1.83	1.66	-
Nonspit					-	1.84	1.71	-

interaction and in the *t*-tests in the rest of the table. Values of *F* and of *t* are typically larger than in the previous analyses.

In summary, the ELE emerged in both models presented with the pictogram input lexicon, but with a significant interaction between model type and presentation condition, confirming our claim that two sources are responsible for the effect.

4 Discussion and Conclusions

We began with the phenomenon, demonstrated from a range of experimental paradigms, that the representation of the exterior letters of words seems to be prioritized in reading—the ELE. We described the precise vertical split-

Table 9: Results of the Analyses of Variance and *t*-Tests for the Split and Nonsplit Networks Presented with the 60 "Pictogram" Input Lexicon.

Input Nodes	Interaction Between Model and Presentation Condition	*t*-Test for Exterior Letters, Comparing Split and Nonsplit Models	*t*-Test for Split Model, Comparing Exterior and Interior Letters	*t*-Test for Interior Letters, Comparing Split and Nonsplit Models	*t*-Test for Nonsplit Model, Comparing Exterior and Interior Letters
1234	**,***	n.s.	***	n.s.	***
2345	***,***	***	***	n.s.	***
3456	***,***	***	***	n.s.	***
4567	***,***	***	***	n.s.	***
5678	***,***	$p = .053^{\dagger}$	***	n.s.	***

Notes: $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.
The *p* value marked with † refers to an effect that is in the unpredicted direction: the exterior letters accrue higher MSE in the split than in the nonsplit model.

ting of the human fovea, an underrecognized fact within the field of visual word recognition, and we hypothesized that part of the explanation for the ELE might lie in the divided nature of the processor. Our simulations used a shift-invariant identity mapping in a split and a nonsplit model. In this approach, letter location is not physically encoded within the initial architecture (there is complete connectivity between input and hidden layers) but is expressed in terms of co-occurring information. In the split model, the most important coding of location concerned the half of the model in which particular pieces of information occur. Our simulations have shown two distinct contributions toward the ELE: one arising within the nonsplit model and one reflecting the vertical division in the split model. Below, we explore related explanations of these two contributions and assess some of the psychological implications.

First, we consider the ELE as it appears in the split model. The data in Tables 1, 4, and 7 suggest a resource-based explanation for the ELE in the split model. Each table shows the MSE for both of the exterior letters for each presentation position of the word. Error is smallest for the left-most letter position when it appears at input node 4 and for the right-most letter position when it appears at input node 5 (in Table 1, the MSE is 0.20 and 0.19, respectively). In these presentations, a single, exterior letter is stranded alone in one half of the model, allowing that letter access to all of the processing resources of that half of the model. In contrast, the remaining three letters must share the resources of the other half of the model. Consider the left-most letter in the word as it appears in input positions 1, 2, 3, and 4. The MSE progressively falls (from 0.68 to 0.20 in Table 1, for instance).

This fall reflects the fact that the letter is being processed in the same half of the model and, for each presentation, has a progressively larger share of the processing resources. The MSE increases sharply when the midpoint is crossed (it is 0.97 for input position 5 in Table 1), reflecting the fact that the letter is now being processed in a different half of the model, using resources that must also represent the rest of the word. The three sets of simulations all demonstrate this precise pattern, for both ends of the presented word in the split model. In contrast to the exterior letters, the interior letters are divided more evenly between the two halves of the split model, and the resulting MSE data reflect the fact that the responsibility of each half of the model for each interior letter is not so skewed as it is for the exterior letters. We will refer to this explanation as the hemispheric division of labor account. Further evidence for this resource-based account is found in the way that the split model learns words in the five different presentation positions shown in Figure 2. For 10 different versions of the split model, we plotted the fall in total MSE for all of the words for each of the presentation positions every 10 epochs of training. We carried out this procedure for the three different lexica. For each lexicon, the central presentation position (input nodes 3, 4, 5, and 6) typically began to show a smaller MSE than the other positions very early in training and maintained this advantage throughout. The MSE for the two single-hemifield presentations typically fell the slowest of all, with the remaining two positions usually accruing intermediate levels of MSE. This collaboration of the two halves of the model resembles a simple example of superadditivity—a feature of human performance on a range of tasks using the two visual hemifields, in which the combined activity of the two hemispheres is apparently greater than the "sum of their parts" (see, e.g., Banich & Belger, 1990).

The second contribution toward the ELE appears to come from the shift-invariant identity mapping itself, given that the nonsplit model produces the ELE in its own right. There is complete connectivity between the input and hidden layers, meaning that exterior letters are defined only by appearances of the stimuli in other locations in the input field. As Table 5 illustrates, the MSE associated with interior letter positions is typically greater for more central input locations; the shift-invariant mapping requires these nodes to support a greater variety of representations, thereby increasing the problems of superpositional storage for these positions. We will refer to this explanation as the superpositional storage account. Support for this account comes from recording the fall in MSE in the nonsplit model for the five presentation positions and for the three lexica, as described above. From very early in training, the central presentation position typically produced the highest MSE of all five positions. The U-shaped curve for MSE across the five positions that was observed for the split model, as described above, was inverted for the nonsplit model; in the nonsplit model, the single hemifield presentation positions typically generated the smallest levels of

MSE throughout training. The mechanisms behind the hemispheric division of labor account and the superpositional storage account may operate additively in the split model.

One of the potential explanations of the ELE we considered involved the split model's disproportionately disrupting the immediate contexts of interior letters more than those of exterior letters. This potential account of the ELE in the simulation with real English words received no support from the subsequent simulations with the random-letters lexicon and the pictogram lexicon. In these two sets of simulations, there is less and less potential to define the constituents of words in terms of their contexts, compared with the real English words, yet the ELE becomes progressively stronger across the three sets of simulations. We therefore reject this potential account based on disrupted contexts; the hemispheric division of labor account and the superpositional storage account remain the principal explanations of the observed behaviors in the three sets of simulations we have reported.

The increase in the size of the ELE from the English words, to the random-letter words, to the pictograms is a marked feature of the data we have presented. This change reflects the increasing difficulty of the task as less and less intermediate structure (i.e. orthotactic constraints and a letter-level of representation) is available to the model. This increased task difficulty was reflected in the number of epochs required by the models to train to criterion for each input lexicon. The longer and more difficult the learning task, the greater is the opportunity for the two sources of the ELE to take effect. The stronger ELE may be seen in increasing significance levels across F and *t* in Tables 3, 6, and 9 and in the appearance of significant effects in some of the single-hemifield presentations to the split model. The learning phase is essential to the network's performance, and so we would predict that presenting novel stimuli to the network would not show the ELE so clearly. The reverse effect to the ELE is observed in human data with the presentation of letter-like nonsense characters (Hammond & Green, 1982; Mason, 1982; Mason & Katz, 1976), and whether this would emerge from our model is a topic for future research.

We turn now to the psychological implications of the modeling results. The models and the task that we have presented are simple and small; only the vertical splitting of one of the models is directly psychologically realistic. What conclusions can be drawn for psychological models of visual lexical processing? The hemispheric division of labor account of the ELE has a direct psychological interpretation, as follows. Any word can be fixated at any point along its length, causing it to be cleanly split in its projection to the two hemispheres. Processing that is directly relevant to word recognition occurs intrahemispherically before the two domains of information are completely merged. Over a history of a variety of fixation points for any one word, the right hemisphere develops an increasingly exclusive relationship with the information that is closer to the left end of the word, and the left hemisphere behaves similarly with respect to the information closer to the right end of the word. This pattern of processing produces the ELE. This account of the ELE has been derived here solely from computational modeling. In retrospect, it might equally well have been produced from a consideration of the facts of foveal splitting and fixation data during reading, in which case the modeling above would be a computational implementation of the argument.

It is less easy to claim that the superpositional storage account has a comparable psychological reality. In this account, we claimed that the weighted connections between the middle region of the input and the hidden layer (in the nonsplit model) came under pressure to represent a greater variety of letters, and so better, less errorful representations emerged toward the edges of the input field, where the exterior letters tend to fall. The two accounts—superpositional storage and hemispheric division of labor—are both cashed out in the same terms in the simple connectionist models we have employed: a better mapping is possible if a greater number of weighted connections is exclusively available to mediate it. In the superpositional storage account, the exclusivity arises from the relatively smaller density of mappings required from the outer regions of the input field. In the hemispheric division of labor account, the exclusivity arises from different parts of the input falling in different halves of the processor. The hemispheric division of labor account may be plausibly applied to real single-word reading; the earlier stages of single-word reading are demonstrably divided between the two hemispheres. The superpositional storage account requires more specific, and less plausible, assumptions about word recognition in the brain; it requires us to believe that the representational substrate for word recognition is coded in terms of absolute location, with specific letter positions possessing dedicated resources. In contrast, the hemispheric division of labor account need only specify location in terms of being in one or other hemisphere. We therefore claim that of the two sources we have identified for the ELE in our simulations, the hemispheric division of labor account has more psychological reality. Our goal has been to model the psychological data, and we therefore conclude that these data are in part explained by foveal splitting and the initial projection of the two parts of a word to different hemispheres. We have directly modeled the ELE in the whole-word priming experiments reported by McCusker et al. (1981); our simulations with different lexica suggest that the effect, stemming from the split nature of the architecture and the nature of the task, is sufficiently strong not to be eclipsed by the orthotactic structure present in the real-world materials.

Our model is a very abstract characterization of word recognition, intended only to represent the psychological reality of foveal splitting. Its input consisted of only four-letter words, reflecting the fact that the relevant human experimentation predominantly uses words of this length. It is likely that each hemisphere has available (from subcortical routes) relatively crude, low-spatial-frequency visual information from the ipsilateral hemifield, thereby providing information about word length (see Sergent, 1987; Corballis & Sergent, 1988). Restricting the model to words of one length therefore has some psychological justification.

Jordan (1990) lists some of the letter groups that have been proposed as processing units in the reading of isolated words: digrams, trigrams, prefixes, suffixes, morphemes, phonological units, and basic orthographic syllabic structures. He proposes exterior letters as just such a perceptual unit. We have developed Jordan's proposal in the context of a psychologically more realistic architecture: the splitting of the visual projection is fundamental to the psychological architecture underlying word recognition. The putative perceptual units listed by Jordan are generated on the strength of emergent formal and statistical structure. Prefixes, for example, are morphological units and also occur sufficiently frequently to have perhaps assumed the status of single entities. The perceptual units we have described are qualitatively different from the other units listed, in that they have their origin in the anatomical structure of the processor. In the modifications that Smith, Jordan, and Sharma (1991) make to the IAM, the length of the word is independently represented, thereby encoding the status of the exterior letters as a qualitatively different type of information. In our own model, the length of the word is not separately represented.

The strongest claim we can make is that the ELE found in reading is solely due to foveal splitting, as we have modeled the phenomenon. A more realistic conclusion may be that a variety of factors conspire to make the prioritizing of the exterior letters of words an adequate and natural strategy for accomplishing word recognition. We consider them in turn.

First, there may be a (small) contribution to the ELE from psychophysical factors. It may also be that lateral inhibition between adjacent letters exists at the abstract, cognitive levels of lexical processing, favoring outside letters.

Second, exterior letters possess a unique informational status, which we have explored elsewhere (Shillcock et al., in press): given a fixation in the middle of a word or just left of the middle, letter location information need only specify which hemifield a letter appears in to define uniquely most words in the English lexicon. If the left hemifield contains *a*, *c*, and *r* and the right hemifield contains *e*, *p*, and *t*, then the word must be *carpet*. This hemifield information about letter location is less adequate for shorter words,[5] and it is the identity of the exterior letters that resolves all of the ambiguities in the three- and four-letter words and virtually all of the remaining ambiguity in the five-letter and longer words. In summary, if letter identities are known, then letter location information can be surprisingly crude, but with exterior letters playing a unique informational role. (Both of these

[5] Although note that less than 5% of four-letter words are ambiguous, like *time* and *item*, which share the same letters in the right and left hemifield following a central split.

factors are the direct result of foveal splitting, which automatically assigns the two parts of the fixated word to different hemispheres and instantiates a "two-slot" processor.)

The third factor, which chiefly concerns monosyllabic words, is that the exterior letters usually represent the onset and coda, and therefore tend to be more informative than the interior letters, better restricting the identity of the word involved. This third factor is relatively fortuitous.

Thus, although our hemispheric division of labor account may be central to the explanation of the ELE, different factors converge to make exterior letters important and to encourage the processor to prioritize them. Indeed, within these constraints, the range of possible solutions to the problem of visual word recognition is much more limited than would first appear. At a more general level we have shown, along with Reggia, Goodall, and Shkuro (1998), for instance, that small-scale neural network models of hemispheric coordination can produce behavior that illuminates processing problems that are solved with the full resources of the brain. The analysis of such models can provide a richer conceptual vocabulary for exploring hemispheric interaction than has emerged from classical models of cognition.

Appendix: Training Set Words

also	away	back	both
call	come	down	each
even	fact	feel	find
from	give	good	hand
have	here	into	just
keep	know	last	late
life	like	look	make
many	mean	more	much
must	only	over	part
same	seem	some	such
take	tell	than	that
them	then	they	this
time	turn	very	want
well	what	when	will
with	work	year	your

Acknowledgments

This research was carried out partly under MRC (U.K.) grant G9421014N, and ESRC (U.K.) research studentship R00429634206. We acknowledge the very helpful comments of Bruce Graham, Gary Cottrell, and an anonymous reviewer. All remaining errors are our own.

Correspondence concerning this article should be addressed to Richard Shillcock, Department of Psychology, 7 George Square, Edinburgh, EH8 9JZ, or to the authors at rcs@cogsci.ed.ac.uk or pmon@cogsci.ed.ac.uk, respectively.

References

Baayen, R. H., Pipenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Banich, M. T., & Belger, A. (1990). Interhemispheric interaction: How do the hemispheres divide and conquer a task? *Cortex, 26,* 77–94.

Brysbaert, M. (1994). Interhemispheric transfer and the processing of foveally presented stimuli. *Behavioural Brain Research, 64*, 151–161.

Butler, B. E., & Merikle, P. M. (1973). Selective masking and processing strategy. *Quarterly Journal of Experimental Psychology, 25,* 542–548.

Corballis, M. C., & Sergent, J. (1988). Imagery in a commissurotomized patient. *Neuropsychologia, 33*, 937–959.

Eriksen, C. W., & Rohrbaugh, J. (1970). Visual masking in multielement displays. *Journal of Experimental Psychology, 83*, 147–154.

Estes, W. K., Allmeyer, D. H., & Reder, S. M. (1976). Serial position functions of letter identification at brief and extended exposure durations. *Perception and Psychophysics, 19*, 1–15.

Fendrich, R., & Gazzaniga, M. S. (1989). Evidence of foveal splitting in a commissurotomy patient. *Neuropsychologia, 27*, 273–281.

Fendrich, R., Wessinger, C. M., & Gazzaniga, M. S. (1996). Nasotemporal overlap at the retinal vertical meridian—investigations with a callosotomy patient. *Neuropsychologia, 34*, 637–646.

Forster, K. I. (1976). Accessing the mental lexicon. In E. C. J. Walker & R. J. Wales (Eds.), *New approaches to language mechanisms*. Amsterdam: North Holland.

Forster, K. I., & Gartlan, G. (1975). *Hash coding and search processes in lexical access*. Paper presented at the Second Experimental Psychology Conference, University of Sydney.

Hammond, E. J., & Green, D. W. (1982). Detecting targets in letter and nonletter arrays. *Canadian Journal of Psychology, 36*, 67–82.

Jordan, T. R. (1990). Presenting words without interior letters: Superiority over single letters and influence of postmark boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 893–909.

Jordan, T. R. (1995). Perceiving exterior letters of words: Differential influences of letter-fragment and non-letter-fragment masks. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 512–530.

Mason, M. (1982). Recognition time for letters and non-letters: Effects of serial position, array size, and processing order. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 724–738.

Mason, M., & Katz, L. (1976). Visual processing of nonlinguistic strings: Redundancy effects and reading ability. *Journal of Experimental Psychology: General, 105*, 338–348.

McCusker, L. X., Gough, P. B., & Bias, R. G. (1981). Word recognition inside out and outside in. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 538–551.

Merikle, P. M. (1974). Selective backward masking with an unpredictable mask. *Journal of Experimental Psychology, 103*, 589–591.

Merikle, P. M., & Coltheart, M. (1972). Selective forward masking. *Canadian Journal of Psychology, 26*, 296–302.

Merikle, P. M., Coltheart, M., & Lowe, D. G. (1971). On the selective effects of a pattern masking stimulus. *Canadian Journal of Psychology, 25*, 264–279.

Mozer, M. C. (1987). Early parallel processing in reading: A connectionist approach. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 83–104). Hillsdle, NJ: Erlbaum.

O'Reilly, R. C., Dawson, C. K., & McClelland, J. L. (1995). PDP++ neural network simulation software. Pittsburgh, PA: Carnegie Mellon University.

Plaut, D. C., & McClelland, J. L. (1993). Generalization with componential attractors: Word and nonword reading in an attractor network. In *Proceedings of the 15th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Plaut, D. C., & Shallice, T. (1994). *Connectionist modelling in cognitive neuropsychology: A case study*. Hillsdale, NJ: Erlbaum.

Reggia, J., Goodall, S., & Shkuro, Y. (1998). Computational studies of lateralization of phoneme sequence generation. *Neural Computation, 10*, 1277–1297.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.

Rumelhart, D. E., & McClelland, J. L. (1981). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89*, 60–94.

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

Sergent, J. (1987). A new look at the human split brain. *Brain, 110*, 1375–1392.

Shillcock, R., Ellison, T. M., & Monaghan, P. (in press). Eye-fixation behaviour, lexical storage and visual word recognition in a split processing model. *Psychological Review*.

Shillcock, R. C., Kelly, M. L., Buntin, L., & Patterson, D. (1997). *Similarities and differences between content and function words: The role of lexical neighbourhoods in the clarity gating task* (Rep. No. EOPL-97-4). Edinburgh: Department of Linguistics, University of Edinburgh.

Shillcock, R. C., & Monaghan, P. (1998). Using anatomical information to enrich the connectionist modelling of normal and impaired visual word recognition. In S. J. Derry & M. A. Gernsbacher (Eds.), *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society* (pp. 945–950). Hillsdale, NJ: Erlbaum.

Shillcock, R., Monaghan, P., & Ellison, T. M. (1999). The SPLIT model of visual word recognition: Complementary connectionist and statistical cognitive modelling. In D. Heinke, G. W. Humphreys & A. Olson (Eds.) *Connectionist models in cognitive neuroscience: The 5th Neural Computation and Psychology Workshop* (pp. 3–12). London: Springer-Verlag.

Smith, P. T., Jordan, T. R., & Sharma, D. (1991). A connectionist model of visual-word recognition that accounts for interactions between mask size and word length. *Psychological Research, 53*, 80–87.

Sperry, R. W. (1968) Apposition of visual half-fields after section of neocortical commissures. *Anatomical Record, 160*, 498–499.

Sugishita, M., Hamilton, C. R., Sakuma, I., & Hemmi, I. (1994). Hemispheric representation of the central retina of commissurotomized subjects. *Neuropsychologia, 32*, 399–415.

Yannakoudakis, E. J., & Hutton, P. J. (1992). An assessment of *N*-phoneme statistics in phoneme guessing algorithms which aim to incorporate phonotactic constraints. *Speech Communication, 11*, 581–602.

Received June 14, 1999; accepted July 24, 2000.