The Optimal Behaviour of a Split Model of Word Recognition Resembles Observed Fixation Behaviour

Richard Shillcock (rcs@cogsci.ed.ac.uk)
T. Mark Ellison (marke@cogsci.ed.ac.uk)
Padraic Monaghan (pmon@cogsci.ed.ac.uk)
Institute for Adaptive and Neural Computation,
Division of Informatics, University of Edinburgh,
2 Buccleuch Place, Edinburgh, EH8 9LW, U.K.

Abstract

We expand upon the case for believing that the initial precise splitting of the foveal projection to the visual cortex fundamentally conditions the whole process of visual word recognition. We explore the optimal behaviour of a split architecture that attempts to divide its processing load equally between its two halves. We successfully model three aspects of fixation behaviour in human readers: (a) the positioning of the optimal viewing position to the left of the midpoint of the word, (b) a displaced Gaussian curve of letter-report accuracy resembling an RVF advantage, (c) the tendency for shorter words not to be directly fixated.

Modelling visual word recognition

The connectionist modelling of visual word recognition in the tradition of the Seidenberg and McClelland developmental model (1989) has led to a richer and deeper understanding of both normal and impaired visual word recognition over the last decade. However, such modelling has not consistently addressed issues of fixation behaviour in reading, for the reason that this tradition of modelling has essentially conceived of word recognition as the abstract mapping between orthographic, phonological and semantic representations, with all of the orthographic input being simultaneously and completely available to the model when any particular word is presented for processing. As with much other research in visual word recognition, this research begins with the abstract problem of distinguishing a target word from 50,000 other words in the lexicon, and the problem is not constrained by what is known about the human visual system.

In contrast, we have argued that the cognitive modelling of word recognition may be advanced by beginning the modelling at an earlier, anatomically constrained stage – that of the precise splitting of a word when it is initially presented to the fovea (Shillcock & Monaghan, 1998). We have demonstrated that "vertically" split, small-scale feedforward networks exhibit behaviour relevant to word recognition when such networks are required to co-ordinate input that straddles the split in the architecture. For instance, such a model automatically prioritises the processing of the end-letters of words (cf. Jordan, 1990), the model's learning and final performance exhibit superadditivity in regard to the separate and combined functioning of the two halves of the model (cf. Banich & Belger, 1990), and its learning evidences a bilateral effect when the same word is presented to each hemifield (see Mohr, Pulvermüller & Zaidel, 1994) (For further discussion, see Shillcock & Monaghan, *in press*; Shillcock & Monaghan, *submitted*). Below we rehearse some of the arguments in favour of this claimed role for foveal splitting in word recognition and we then compare the optimal behaviour of such a model with the observed behaviours of human readers.

A split fovea conditions word recognition

Research with commissurotomy patients reveals that the human fovea is precisely vertically split (Fendrich & Gazzaniga, 1989; Fendrich, Wessinger & Gazzaniga, 1996; Sugishita, Hamilton, Sakuma & Hemmi, 1994). When a word is fixated the part of the word in the left visual field (LVF) goes initially to the right hemisphere (RH) and the part in the right visual field (RVF) goes to the left hemisphere (LH). After this initial projection we claim that substantial processing directly relevant to word recognition occurs *intra*hemispherically before any putative complete sharing of information from the two hemifields. The argument for this claim can only be summarised here, but includes the following points.

(a) There are robust hemispheric differences in word recognition, which would not be apparent if all word recognition occurred after complete sharing of information (Hellige, 1995; Mohr, Pulvermüller & Zaidel, 1994).

(b) Cortical processing involves orchestrations of activity, replete with recurrent connectivity, rather than unidirectional cascades of processing towards a single abstract goal (see, e.g., VanEssen & Felleman, 1991). Activity in the (neatly split) primary visual cortex seems to be crucial for visual awareness (see, e.g., Cowey & Stoerig, 1992).

(c) Cells in the visual cortex require their receptive fields to contain the vertical midline in order to have a direct callosal connection (Whitteridge, 1965).

(d) Even very high level lexical processing (reading, oral spelling, mental rotation of words) is frequently impaired by unilateral neglect in a predictable spatial manner (Hillis & Caramazza, 1990).

(e) There is direct connectivity between the RH's inferior-temporal cortex and the LH's language regions (DiVirgilio & Clarke, 1998).

(f) The word-beginning superiority effect interacts with hemispheric dominance for language for foveally presented stimuli (Brysbaert, 1994). In particular, see Brysbaert (1994) for further discussion of the issue of foveal splitting.

The evidence cited above indicates that researchers in visual word recognition should explore the cognitive consequences of foveal splitting. As we will demonstrate below, the parsimony of the cognitive modelling accounts based on a generic "Split Model" bears out the claim that foveal splitting fundamentally conditions even the high-level cognitive aspects of visual word recognition.

A split word is already informative

Let us begin by assuming that a word is fixated somewhere close to its midpoint. We have seen that the two halves of the word are then projected to different hemispheres. Conventional modelling of word recognition assumes two partly contradictory next steps: first, that the information should all be moved into one hemisphere, and second, that detailed information is represented about the location of all the individual letters. We can assess these two proposals by considering just how much information is in our split starting position, which in effect resembles a model of word recognition with just two letter locations, the RVF and the LVF. If each word is split as close to its midpoint as possible, then *carpet* is split as [*a*, *c*, *r*] and [*e*, *p*, *t*], ignoring the precise order of the letters. This already gives us enough letter and location information to identity *carpet* uniquely. In contrast, [*i*, *t*] and [*e*, *m*] is ambiguous between *item* and *time*, and we need more letter-location information to distinguish between these two possibilities. What proportion of the lexicon is unambiguously specified, like *carpet*, by such a two-slot model?

The answer is that 98.6% of English words in the CELEX database (Baayen, Pipenbrock & Gulikers, 1995) are like *carpet*. This result is intuitively plausible for longer words, so let us consider four-letter words only; four-letter words are the most labile case in this analysis. When no information is given about letter position in four-letter words, then 34% are ambiguous (i.e. they are anagrams of other words). When four-letter words are split, the RVF *versus* LVF information disambiguates all but 4.7% of words (like *item*). These data constitute an empirical result. If the processor can fixate near the mid-point of a word, the split information is already very informative of word identity. A split word is an attractive goal rather than an inconvenience caused by the anatomy and something to be transcended.

An ambiguity level of 4.7% is still not ideal and more information is required. Happily this information is readily available from an aspect of split architectures that we have described elsewhere (Shillcock & Monaghan, 1998; Shillcock & Monaghan, *submitted*): split architectures naturally prioritise the processing of the outside letters of words when they are trying to co-ordinate inputs that are typically spread across the two halves of the model. Identifying the outer letters disambiguates all of the four-letter words and leaves the greatest ambiguity at five-letter words, of which only 0.62% remain ambiguous like *trial* and *trail*.

Thus, we see that splitting words is an informationally attractive precursor to identification, and that fast normal visual word recognition may not need exhaustive information about letter location. We now consider the further processing implications of split architectures by exploring their optimal behaviour.

Optimal behaviour of a split model

One of the factors affecting the behaviour of a split network concerns the balancing of the independent activity in its two halves. *Ceteris paribus*, such a network might be expected to function optimally if labour is evenly balanced between its two halves. A simple approximation to this state of affairs would emerge from fixating each word at its physical midpoint; each half of the model would have to cope with the same quantity of orthographic input We term this fixation strategy the *middle-justified strategy*. Alternatively. the model could arrange to fixate each word so that the left end of each word appears at the left end of a putative foveal window. This *left-justified strategy* is actually suggested by Legge et al. (1997) as an approximation of the behaviour of their model of fixation behaviour. A further (hypothetical) model might be a *right-justified strategy* .

Let us assume complete control over word fixation, so that each word is individually fixated at the optimal location (Although this strategy presupposes prior knowledge of the word's identity, such an analysis gives us an illuminating picture of the optimal behaviour of a split model.) In this case, the optimal fixation point for each word in the lexicon will be one that creates the minimum competition in the right and in the left half of the model. We developed a splitting algorithm that iterated through all the words in the lexicon, in random order each time, moving the split-point (which equals the fixation point) so as to minimise $1/lr$, in which l and r are respectively the numbers of occurrences of the left and right half of the word in the left and right halves of the split words in the lexicon[1]. The algorithm begins with a random placement of the split-point in every word, and converges after many iterations on one of a number of qualitatively closely comparable outcomes. Figure 1 shows the mean of ten different runs of the algorithm.

Each line of the graph shows the distribution of split-points for words of a particular length. The graph shows that the mean split-point is to the left of the physical midpoint of the word, and that this split-point predominates for the longer words. (We discuss the shorter words below.) This result matches the observation that the starts of English words are typically more informative than their ends. It confirms that a principle of dividing the processing equally between the two halves of the model produces a psychologically realistic outcome for English: word recognition is better for most words when fixated to the left of the midpoint, at an *optimal viewing position* (OVP) (O'Regan, 1990). One heuristic that approximates this optimal behaviour is to fixate words slightly to the left of their midpoints. The orthodox interpretation of the OVP is that it is the result of the positioning of a limited, high resolution foveal window over the most informative part of the word. Our results demonstrate that the OVP may equally be interpreted as a split processor attempting to divide the processing load evenly between its two halves.

The splitting algorithm produces detailed behaviour that further approximates the observed performance of human readers. The algorithm is free to choose a fixation point adjacent to any letter in the current word, the choice being determined solely by the reduction of $1/lr$. On some occasions the algorithm places the split-point to the left of the first letter or to the right of the last letter of the word: ||*rip* or *the*||. In such a case, the algorithm has placed the

---

[1] Both l and r can only fall to 1, representing their occurrence in the current word.

word exclusively in one hemifield, balancing the uninformativeness of an empty hemifield against the increased informativeness of, for instance, three letters rather than two in the other hemifield. As Figure 2 shows, this aspect of the optimal behaviour of the split model resembles the observed human failure to fixate short words, in inverse proportion to word-length (Rayner & McConkie, 1976).

This analysis provides a novel explanation for failure to directly fixate words in text: it is inappropriate to fixate on a word that is so short that its division gives the two halves of the processor only very short, multiply ambiguous parts of words. We propose that this explanation contributes to the observed human behaviour along with other accounts such as the increased predictability of the (predominantly short) function words and the parafoveal preview of short words.

We can see some of the computational implications of the different fixation strategies, introduced above, by arranging every word in the lexicon "vertically", one above the other, in the way suggested by each strategy and then calculating the entropy of the distribution of letters in each

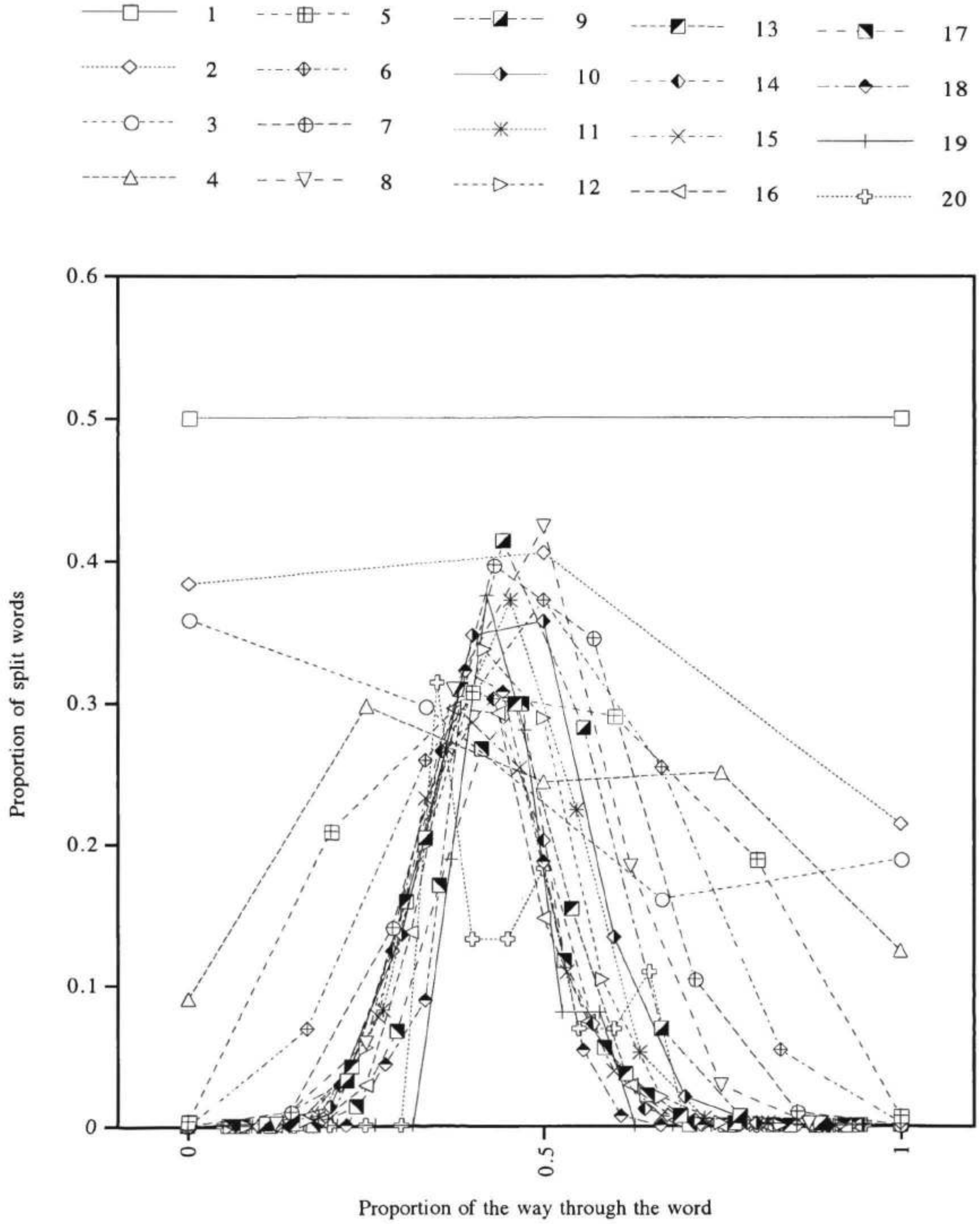

Figure 1: The curves, for each word-length, of the proportion of words optimally split at different proportions of their total length.

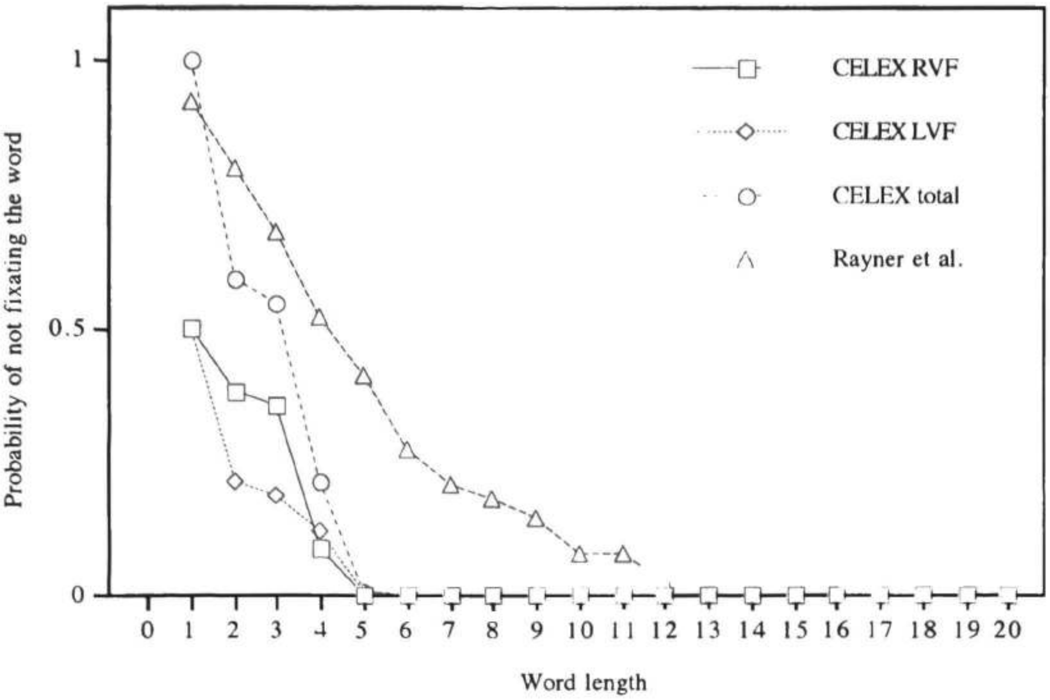

Figure 2. The probability of a word not receiving a direct fixation, as a function of word length. The "CELEX" curves show the probability that the splitting algorithm positions the word completely in the LVF or RVF. The human data are from Rayner and McConkie (1976).

vertically aligned letter location. Thus, for the left-justified strategy, the leftmost column of letters is the first letter of each word in the lexicon, and the entropy of this distribution reflects the skewedness of this distribution. For the middle-justified strategy and for the optimal strategy produced by the splitting algorithm, the fixation point in each word is vertically aligned. We present the results of this analysis in Figure 3. The entropy curves allow us to compare the different fixation strategies. A consistently high curve means that there is high quality information spread across the relevant part of the visual field.

First, note that in Figure 3 the left- and right-justified curves are arbitrarily positioned and may be moved along the *x*-axis to align them with the respective ends of any foveal window. As Figure 3 shows, the middle- and optimally-split curves provide the most symmetrical spread of information across the visual field, although there is no significant difference between the total entropy of the left- and middle-justified strategies. The optimal strategy, as might be expected from the nature of the algorithm, produces the most sustained spread of high quality information across the visual field. Note that the optimal strategy does not need to imply that reliable letter recognition occurs in a window as wide as 26 letters, only that high quality information can be effectively spread across the visual field. Such a spread of information maximises the utility of any parafoveal processing. It naturally emerges from the splitting algorithm that a span of 26 letters is an appropriate span; this span matches observed human behaviour (McConkie & Rayner, 1975). Finally, we might note that the entropy curve from the optimal strategy is asymmetric, in that the curve in the RVF falls off less sharply than that in the LVF. Brysbaert, Vitu and Schroyens (1996) claim that letter-report accuracy (in the form of an extended OVP curve) follows a Gaussian distribution across both foveal and parafoveal parts of the visual field; they claim that the reported RVF advantage for word recognition may be subsumed into this EOVP curve. The entropy curve for the optimal strategy in Figure 3 is in agreement with this EOVP curve: letters are reported at different locations with respect to the fixation point with an accuracy in proportion to the probability that useful information may be found at that position. The optimal-strategy curve in Figure 3 indicates this probability. Thus, the EOVP curve may be accounted for without any necessary role for LH dominance for language.

Discussion

We have shown that a word split between the two hemifields and between the two hemispheres is potentially a very informative starting point for visual word recognition. A midpoint split is a simple fixation strategy but it provides an effective starting point for word recognition. A more sophisticated, optimal split-point for each word illustrates the upper bound for a split processor that is attempting an equal division of labour. This optimal strategy has many features that closely correspond to reading behaviour in human subjects, suggesting that skilled readers have developed fixation behaviour that is perhaps closer to the optimal behaviour than is the simple midpoint split strategy.

One possible way in which this distributional analysis of a full-sized lexicon might be related to the connectionist

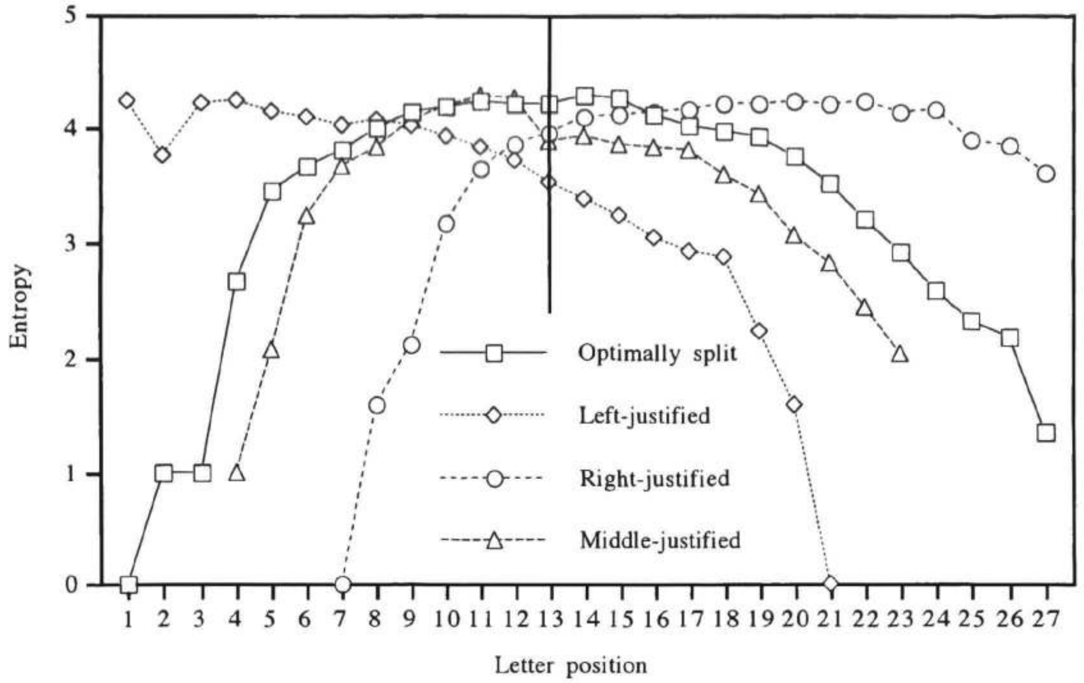

Figure 3: Entropy curves for the optimal splitting strategy, versus other fixation strategies.

modelling with small-scale networks we describe elsewhere is for us to assume that the processor can indeed deal with any word presented at any position relative to the fixation point, but that for each word type there is a strongest, best template situated in one particular position, contingently defined by the rest of the words that must be stored.

Conclusions

We have shown that if we follow through the implications of the initial precise splitting of the foveal projection, to consider the effects on the later stages of visual word recognition, then psychologically realistic behaviours result. These behaviours emerge naturally from the proposed Split Model. We have provided accounts for a number of different phenomena in visual word recognition, for which alternative yet separate accounts previously existed. The coherence of the accounts based on foveal splitting, and their grounding in known anatomy are arguments for favouring these accounts of the phenomena and for further exploring the claim that foveal splitting fundamentally determines visual word recognition at all levels.

References

Baayen, R.H., Pipenbrock, R. & Gulikers, L. (1995). The CELEX Lexical Database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

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.

Brysbaert, M., Vitu, F. & Schroyens, W. (1996). The right visual field advantage and the optimal viewing position effect: On the relation between foveal and parafoveal word recognition. *Neuropsychology*, **10**, 385– 395.

Cowey, A. & Stoerig, P. (1992). Reflections on blindsight. In *The neuropsychology of consciousness* (Eds. A.D. Milner & M.D. Rugg), Academic Press, London.

DiVirgilio, G. & Clarke, S. (1997). Direct interhemispheric visual input to human speech areas. *Human Brain Mapping*, *5*, 347–354.

Fendrich, R. & Gazzaniga, M.S. (1989). Evidence of foveal splitting in a commissurotomy patient. *Neuropsychologia*, Vol. 27, No. 3, pp. 273-281.

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

Hellige (1995). Coordinating the different processing biases of the left and right cerebral hemispheres. In F. Kitterle, (Ed.) *Hemispheric communication: Mechanisms and models*. Law. Erl. Assoc., Hove.

Hillis, A.E. & Caramazza, A. (1990). The effects of attentional deficits on reading and spelling. In A. Caramazza (Ed.) *Cognitive Neuropsychology and Neurolinguistics: Advances in Models of Cognitive Function and Impairment*. Law. Erl. Assoc., Hove, UK.

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.

Legge, G.E., Klitz, T.S. & Tjan, B.S. (1997). Mr. Chips: An ideal-observer model of reading. *Psychological Review*, *104*, 524–553.

McConkie, G.W. & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, **17**, 578–586

Mohr, B., Pulvermüller, F. & Zaidel, E. (1994). Lexical decision after left, right and bilateral presentation of function words, content words and non-words: Evidence for interhemispheric interaction. *Neuropsychologia*, *32*, 105–124.

O'Regan, J.K. (1990). Eye movements and reading. In E. Kowler (Ed.) *Eye movements and their role in visual and cognitive processes*, North Holland: Elsevier Science Pub.

Rayner, K. & McConkie, G.W. (1976). What guides a reader's eye movements? *Vision Research*, **16**, 829–837.

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

Shillcock, R.C. & Monaghan, P. (*submitted*). The computational exploration of visual word recognition in a split model.

Shillcock, R.C. & Monaghan, P. (1998). Using anatomical information to enrich the connectionist modelling of normal and impaired visual word recognition. *Proceedings of the 1998 Cognitive Science Society Conference*, Wisconsin, 945–950.

Shillcock, R., Monaghan, P. & Ellison, T.M. (*in press*). The SPLIT model of visual word recognition: Complementary connectionist and statistical cognitive modelling. Proceedings of the 5th Neurocomputation and Psychology Workshop (eds. G. Humphreys et al.), Springer-Verlag.

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

VanEssen, D.C. & Felleman, D.J. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, **1**, 1–47.

Whitteridge, D. (1965). Area 18 and the vertical meridian of the visual field. In (ed. E.G. Ettlinger) *Functions of the Corpus Callosum.*, pp. 115–120. Churchill, London.