

Supporting the Serial in the SERIOL Model

Carol Whitney

Department of Linguistics
University of Maryland
College Park, MD 20742
USA
301-384-3391
cwhitney@cs.umd.edu

Short Title: Serial Lexical Access

Abstract

It is commonly assumed that orthographical lexical access in visual word recognition takes place in parallel, with all letters activated at the same time. In contrast, in the SERIOL model of letter-position encoding, letters fire sequentially (Whitney, 2001). I present further support for such seriality on several fronts. (1) The reasons that led to the rejection of serial encoding are shown to be invalid, and the virtues of a serial encoding are discussed. (2) The SERIOL model's serial mechanisms provide a natural account of counterintuitive letter-perceptibility patterns, and correctly predict the temporal evolution of these patterns. (3) Via a simulation, I show that serial lexical access accounts for conflicting data which indicate both a presence and absence of positional effects at the lexical level. In contrast, both the perceptibility and lexical results are inexplicable under the assumption of parallel lexical access.

Acknowledgments

I thank Jonathan Grainger for sharing his data, and for stimulating conversations. I also thank Gordon Legge, Stephen Mansfield, and Susana Chung for sharing their data, with appreciation to Stephen Mansfield for preparing that data.

Introduction

As the brain performs visual word recognition, it cannot merely encode letter identities, due to the existence of anagrams. Rather, the arrangement of the letters must also be represented. Masked form-priming studies have indicated that this encoding is based on the order of letters (relative position), not their absolute string positions (Humphreys, Evett & Quinlan, 1990; Perea & Lupker, 2003; Peressotti & Grainger, 1999; Schoonbaert & Grainger, 2004). For example, the prime BLCN facilitates lexical-decision responses to the French word BALCON, but the prime BCLN does not. Furthermore, the prime B LC N, in which the string position of the target letters is preserved, does not provide any more facilitation than BLCN (Peressotti & Grainger, 1999). Based on such data, I have proposed, in the SERIOL model of letter-position encoding, that *non-contiguous bigram* representations contact the lexical level (Whitney & Berndt, 1999; Whitney, 2001). Such units represent an ordered pair of letters that are not necessarily next to each other. For example, the word CART would activate bigrams CA, AR, and RT, as well as CR, CT, and AT. Grainger later also endorsed such units, using the more commodious moniker *open bigrams* (Grainger & Whitney, 2004; Schoonbaert & Grainger, 2004), which I will adopt.

But how are open bigrams activated? It is commonly assumed that complex, location-invariant representations are formed in parallel via increasing receptive-field sizes. Some have suggested that multi-letter units are activated in this way (Mozier, 1991; J. Grainger, pers. comm.; S. Dehaene, pers. comm.). In contrast, in the SERIOL model, open bigrams are activated by a serial encoding of letter order (e.g., in our CART example, letter unit C fires, then A, then R, and then T). A bigram unit responds when its constituent letters fire in the proper order. For example, bigram CR is activated when R fires after C, but not when C fires after R.

The serial representation of letter order is the most central and controversial aspect of the SERIOL model. The overall goal of this article is to further support this claim of serial processing. More specifically, I offer such support on three fronts, as follows.

(1) I argue for a serial encoding of letter order in general. I examine reasons against a serial encoding and show that they are not valid. The benefits of a serial encoding are discussed, and direct experimental evidence for a serial readout of the visual image is reviewed.

(2) I argue for the SERIOL mechanisms in particular, based on the capacity of the model to explain some otherwise inexplicable experimental results. For a short string unilaterally presented at a large eccentricity, the letter *farthest* from fixation is the best perceived (i.e., the first letter in the LVF and the last letter in the RVF) (Bouma, 1973; Estes, Allemeyer & Reder, 1976; Legge, Mansfield & Chung, 2001). This counterintuitive result follows naturally from the proposed serial mechanisms. A new analysis of previously presented data (Legge et al., 2001) confirms predictions regarding the temporal evolution of this perceptibility pattern.

(3) The letter and bigram dynamics imply that bigrams are also activated serially. I demonstrate, via a simulation, how the serial activation of bigrams allows the reconciliation of some conflicting experimental data. Recent studies have demonstrated an absence of a positional effect for primes comprised only of letters from the target word (Granier, Grainger, Farioli & van Heuven, 2005). However, a positional effect is observed for primes containing non-target letters (Schoonbaert & Grainger, 2004), and in perceptual and aphasic error patterns (Humphreys, Evett & Quinlan, 1990; Montant, Nazir & Poncet, 1998; Whitney & Berndt, 1999). This simulation replicates both the absence and presence of positional effects. It supersedes a previous simulation, in which positional effects arose from positional differences in bigram activation levels (Whitney & Berndt, 1999). In the present simulation, bigram activations do not vary with position, and positional effects (when present) arise directly from serial bigram activations.

In the following, general arguments for a serial encoding are presented first. Then the SERIOL model is reviewed. Next, discussions of the second and third topics above are presented in turn. Within each of these discussions, I review the relevant experimental results, and then demonstrate how the serial mechanisms of the SERIOL model account for the data.

Discussion of Serial Encoding

First I clarify what is meant by serial encoding. The heart of the matter lies in the activation of letter representations. Are all letters activated at the same time, or are the letters activated in sequence? The latter possibility has two variants. Letters may be activated strictly serially, with only one letter active at a time. That is, the first letter fires; then the first letter stops firing and the second letter fires, etc. The SERIOL model advocates such seriality. Alternatively, a letter may continue firing even after subsequent letters start to fire. That is, the first letter fires; then the second letter starts firing, but the first letter continues to fire, etc. By the time that the final letter fires, all letters are firing. The SOLAR model advocates this type of sequentiality; letter order is encoded by activation level, not directly by order of firing (Davis, 1999). In the following, I will specifically consider strict seriality. However, many of the following arguments are also applicable sequential activation of letters in general.

Another issue involves processing routes. The encoding of letter order not only supports a direct orthographic route to lexical information, it also subserves a sub-lexical, phonological route. While there is some controversy concerning the capabilities of such a sub-lexical route (Coltheart et al., 2001; Harm & Seidenberg, 1999), these leading models of visual word recognition have converged on the assumption that two routes are necessary in order to explain patterns of developmental and acquired dyslexia. Most models assume that the same encoding of letter order subserves both routes. However, Coltheart's dual-route model assumes a parallel encoding for the lexical route, and a serial encoding for the sub-lexical route. In the following, I will concentrate on the lexical route, and use the term *lexical access* to refer to lexical activation via that route. If referring to the sub-lexical route, I will denote it explicitly.

Early accounts of string processing widely assumed that the visual image was read out serially (Gough, 1972; Harcum & Nice, 1975; Mewhort, Merikle & Bryden, 1969; Sperling, 1963). However, it is currently generally assumed that all letters are activated in parallel and that lexical access occurs in

parallel (Coltheart, Rastle, Perry, Langdon & Ziegler, 2001; Harm & Seidenberg, 1999; McClelland & Rumelhart, 1981). First, I discuss why the shift toward the assumption of parallel read out and lexical access occurred, and demonstrate that these reasons are not well motivated. Having considered arguments against a serial encoding, I then discuss evidence and arguments for a serial encoding.

Why not Serial?

It seems that a combination of factors has contributed to the abandonment of the assumption of serial processing along the direct route to the lexicon. I will discuss each topic in turn.

Length Effect

It is generally assumed that the question of serial versus parallel processing can be adjudicated on the basis of reaction times (RTs). If there is no effect of the number of items to be processed, this is taken as evidence of parallel processing. If RTs increase with the number of items, this is taken to reflect serial processing. When lexical-decision experiments were performed in which frequency was well controlled, no effect of string length emerged (Frederiksen & Kroll, 1976; Richardson, 1976). Therefore, these results have been taken as evidence for parallel lexical access.

However, these assumptions are not necessarily warranted. For example, parallel processing of letters could yield a length effect due to inhibitory effects of decreasing acuity with string length. Conversely, serial processing could fail to yield a length effect if an increased number of letters also has a counterbalancing facilitatory effect. For example, after the final letter fires, it may take longer to reach response criterion for shorter words (than for longer words), due to less bottom-up input from fewer letters. Thus if inhibitory and facilitatory effects of more letters were to cancel each other out, there would be no length effect, despite a serial encoding.

A recent study demonstrates the need for these more complex scenarios. New and colleagues (2005) undertook an investigation of the effect of word

length based on the English Lexicon Project, which is a on-line database of lexical-decision RTs for over 40,000 words (Balota, Cortese, Sergent-Marshall & Spieler, 2004). This huge amount of data allows a highly accurate analysis of the influence of different variables. Once the effects of frequency, number of syllables and orthographic-neighborhood size were factored out, they found that RTs actually *decrease* with increasing string length for words of three to five letters¹, are constant with string length for words of five to eight letters, and increase with string length for words of eight or more letters. This same pattern remained for two subsets of the data - monomorphemic nouns and bisyllabic words. Thus string length has differing effects over different lengths.

It is highly unlikely that these effects reflect differences in the method of lexical access. Rather, these results most likely indicate that the effect of length is the sum of multiple influences, where the relative strength of opposing components varies with length. For example, the data are explained by assuming a facilitatory component which decreases in strength as word length increases, coupled with a steady linearly increasing cost of string length due to a serial encoding. Such a facilitatory effect may arise from decreased settling time at the word level due to more bottom-up activation, and this effect may strongest for short words and then taper off, reaching a ceiling level. Thus for short words, the facilitatory effect dominates; for medium-length words, the two components cancel each other out; for long words, the cost of seriality dominates.

Furthermore, we have demonstrated that a length effect can be abolished (Whitney & Lavidor, 2004). It is well known that for unilateral presentation, there is a length effect in the LVF, but not the RVF (Bouma, 1973; Ellis, Young & Anderson, 1988; Young & Ellis, 1985). This difference is present even if the location of the first letter is held constant as string length is

¹It is likely that the reason that this facilitatory effect of word length has not been previously observed is that the effect of orthographic-neighborhood size (N) was not controlled. N is the number of words that can be formed by changing one letter of the target to another letter (Coltheart, Davelaar, Jonasson & Besner, 1977). High N is actually facilitatory for words in lexical decision (Andrews, 1997; New et al., 2005). Because N generally decreases with word length, reduced N facilitation for longer words may have canceled out the facilitatory effect of more letters.

increased, indicating that it is not related to the acuity of the initial letter (Young & Ellis, 1985). Therefore, this asymmetry has been taken to reflect different modes of lexical access in the two hemispheres, with inefficient, serial RH processing and efficient, parallel LH processing (Ellis et al., 1988; Young & Ellis, 1985). However, based on predictions from the SERIOL model, we have succeeded in abolishing the LVF/RH length effect in lexical decision by using positional manipulations of contrast level. Brightening the second and third letters and dimming the sixth letter (if present) of four- to six-letter LVF words abolished the length effect (Whitney & Lavidor, 2004). Such an adjustment could not have converted inefficient, serial processing into parallel processing. Rather, the manipulation likely created a more optimal activation patterns, thereby preferentially decreasing settling time at the word level for longer words, and obliterating the length effect.

In sum, reaction times are not solely a function of how long it takes to activate the letter representations. The above results demonstrate that the presence or absence of a length effect cannot be used to determine whether lexical access occurs serially or in parallel.

Uniqueness Point

Another way of trying to get at the issue of serial versus parallel processing is to consider uniqueness point (UP) - the position at which a word can first be distinguished from all other words. For example, *actress* is the only word starting with *actr*; its UP is 4. For spoken word recognition, it has been shown that words with early UPs are responded to more quickly than words with later UPs (Goodman & Huttenlocher, 1988; Radeau et al., 1989). This is consistent with the cohort model of word recognition, in which lexical access occurs serially in response to the speech signal, and the set of possible words is progressively thinned until only one candidate remains (Marslen-Wilson & Welsh, 1978). Therefore, it has been assumed that if orthographic lexical access is serial, there should also be an effect of UP. An early study showed no effect of UP on visually presented words (Radeau, Morais, Mousty, Saeuens & Bertelson, 1992), leading to a conclusion that letter activations and lexical

access occur in parallel.

However, as Kwanter and Mewhort (1999) pointed out, this study measured uniqueness point phonetically, but uniqueness point should be defined orthographically for visual word recognition. Manipulations of orthographic uniqueness point (OUP) have shown that words with early OUPs are processed faster than words with later OUPs, leading to the conclusion that lexical access in visual word recognition occurs serially from left to right (Kwanter & Mewhort, 1999; Lindell, Nicholls & Castles, 2003).

However, the assumption that the presence or absence of an OUP effect can distinguish between serial and parallel processing should also be approached with caution. The notion that serial processing should necessarily give an OUP effect depends on the problematic assumption that lexical access is terminated when the OUP (or UP) is reached. However, such an assumption incorrectly predicts that *actrem* should be perceived as *actress*. Rather, all letters or phonemes must be processed in case they form a novel word. If all elements are obligatorily processed, serial lexical access without an effect of uniqueness point remains a logical possibility. (It is likely that the observed effects of uniqueness point arise from a competitive advantage for early OUP words which does not directly reflect termination of lexical access at the point of a single lexical candidate.)

Conversely, the assumption that the presence of an OUP effect proves serial processing is also questionable. By definition, a late OUP word will match at least one other word on a large number of letters, while this is not necessarily the case for an early OUP word. As Lamberts (2005) points out, the late OUP words used by Kwanter and Mewhort (1999) did indeed have a greater overlap with other words than the early OUP words. Lamberts argues that their data still support serial lexical access, but letters could be activated in any order, not necessarily from left to right. However, Lamberts' analysis neglects possible interactions at the lexical level. When OUP is confounded with orthographic overlap, parallel letter activation could yield an advantage for early OUP words due to reduced competition from other words. Thus, the presence of an OUP effect may indeed reflect seriality, but doesn't necessarily. Here I am presenting an argument against an analysis that would support

my claim of seriality, as it is important to be clear about what conclusions can be drawn from RT data.

Word Superiority Effect

Using a forced choice task, Reicher (1969) demonstrated that report accuracy is better for a letter within a word than for a single letter. This phenomenon was dubbed the Word Superiority Effect (WSE). Rayner and Pollatsek (1989) argued that the WSE rules out serial lexical access, because letter-by-letter activation would imply that the non-initial letters of a word would take longer to identify than a single letter. Therefore, such letters should be at a disadvantage, rather than at an advantage.

Davis (1999) provides an excellent discussion of the limitations of their argument, reviewed here. Their argument embodies the assumption that once a letter is identified, it can be reported accurately. That is, the difficulty in letter report lies in the initial processing of the letter; this processing would be delayed in a serial model, creating a disadvantage for word stimuli. Although not explicitly pointed out by Davis, the assumption of all-or-nothing identification is clearly false, as form-priming effects occur in the absence of conscious awareness of the prime (Humphreys et al., 1990). That is, letter representations can be activated sufficiently to provide priming, but not enough to allow report.

So an alternative explanation of report difficulty is that the letter representation is not sufficiently activated to drive naming of that letter. Therefore, additional activation via lexical support would stabilize the representation of the letter, increasing report accuracy. This explanation of the WSE is equally compatible with parallel or serial activation of a lexical representation. That is, while it may take longer to activate a non-initial letter in a word (than a single letter), the additional excitation from the lexical level may well override any cost of seriality, creating a WSE.

However, the cost of this seriality may well be evident in the form of decreased accuracy for the final letters. In contrast, a parallel model would predict that all the letters of a word should be approximately equally well

perceived (with perhaps some advantage for the outer letters, due to reduced lateral masking). The data is more consistent with the serial than the parallel prediction. For four-letter words, the first letter is perceived the best of all the letters, and the second letter is perceived better than the third letter (Rumelhart & McClelland, 1982). Thus positional effects are present, consistent with left-to-right activation of letters. (However, perceptibility is not monotonically decreasing across the string: the fourth letter is perceived better than the third letter. I discuss below how such a final-letter advantage could arise in a serial encoding.) Thus serial lexical access is not inconsistent with the WSE, and explains the positional pattern better.

Rate-Coding Models

The bias toward parallel processing was also strengthened by the influence of the Parallel Distributed Processing (PDP) books (Rumelhart & McClelland, 1986). These volumes employed models that were based on the rate-coding hypothesis, wherein information is coded via average firing rate, rather than the precise timing of individual spikes.

In general, the rate-coding hypothesis was motivated by neural-recording experiments which showed that firing patterns were not reproducible across trials (Adrian, 1928). Thus, it was thought that the mechanisms underlying spike timing were inherently noisy, and therefore spike timing could not be used to encode information (Burns, 1968). In contrast, average firing rate was reproducible across trials, and firing rate varied with the strength of the variable being manipulated (Adrian, 1928). Therefore, it was thought that information was encoded in the average firing rate. The PDP volumes were developed in this milieu and employed this assumption.

These volumes had an enormous impact, reviving interest in computational modeling by showing how the back-propagation algorithm could be used to train multi-layer networks with nonlinear activation functions. The assumptions of rate coding and distributed parallel processing came to dominate the field. Furthermore, the success of the Interactive Activation Model (McClelland & Rumelhart, 1981) in explaining the WSE also probably con-

tributed to a parallel-processing bias in this area of research. (Although, as discussed above, facilitation from the word to the letter level could just as well arise under serial lexical access.)

However, recent experiments using more naturalistic stimuli have demonstrated that the assumption of unreliable spike generation is unwarranted. For temporally varying stimuli, spike timing is highly reproducible across trials, at the time scale of a single millisecond (Mainen & Sejnowski, 1995; Rieke, Warland, de Ruyter von Stevenink & Bialek, 1997). For example, recordings of a motion-sensitive neuron in the fly visual system showed high reproducibility of spike trains across trials for a given dynamic velocity pattern (de Ruyter van Stevenink, Lewen, Strong, Koberle & Bialek, 1997). Reproducible spike timing has also been shown in the representation of static visual images in short-term memory (Lee, Simpson, Logothetis & G., 2005). Thus, contrary to earlier assumptions, neural spiking mechanisms are highly precise. Consistent with this finding, recent computational models, including the SERIOL model, have relied on spike timing to provide representational mechanisms (e. g., Hopfield, 1995; Lisman & Idiart, 1995). Moreover, recent theories of perception have focused on the role of spike timing and oscillatory activity to bind and process information (VanRullen & Koch, 2003; Ward, 2003).

Thus one of the underlying motivations for the assumption of parallel processing - that information is represented via rate codes due to unreliability of neural spiking dynamics - proved unwarranted. Yet the bias toward parallel processing, driven in part by the impact of PDP research, remains in the area of visual word recognition. However, the popularity of a certain style of modeling does not provide any hard evidence as to the nature of lexical access.

Why Serial ?

Having refuted arguments *against* serial lexical access, I next discuss evidence and arguments *for* seriality.

Experimental evidence for serial readout

As we have seen, RTs cannot be used to investigate whether a string's letters are processed serially or in parallel. However, seriality can be investigated by using time directly. Harcum and Nice (1975) used this approach in an experiment in which two six-letter pseudoword strings were sequentially presented, and the task was to report the letters perceived. The first string was presented for 70 ms, and the second string was presented for 35 ms. Temporal order significantly interacted with letter position. Subjects tended to report the first letter of the first string, and the second through sixth letters of the second string. Such an interaction was present for nine of the 10 subjects, but the position of the crossover point between the two strings varied with subject.

As these strings were not words, it is unlikely that these results arose from interactions between the two strings at the lexical level. However, even if there were top-down lexical influences, this would not explain the positional interaction with presentation order. Rather, these results provide strong evidence for serial readout of the visual image(s). There was only enough time to process the initial letter(s) of the first string before the stimulus changed to the second string.

Provision of Encoding Mechanisms

A serial encoding accomplishes two things. It provides a way to bind positional information to letter identity, and it allows the formation a location-invariant representation. I will discuss the necessity of each of the functions, and discuss difficulties with non-serial mechanisms.

There is evidence for the existence of *position-independent* letter detectors. Such a detector is activated by any occurrence of that letter, independently of its retinal location or string position. In a series of experiments using consonant trigrams, Peressotti and Grainger (1995) demonstrated that priming could occur across string positions and retinal locations. This indicates that a given letter in different string positions activates the same letter-level representation.

This conclusion is further supported by the existence of letter-position dyslexia in Hebrew subjects suffering from left occipitoparietal lesions. These subjects made reading errors that were characterized by migration errors within a word; that is, errors were predominately anagrams of the target word (Friedmann & Gvion, 2001). This pattern suggests that the encoding of letter identity can be separated from the encoding of position.²

However, there must be some way of binding positional information to such position-independent representations. Serial firing of such letter units meets this requirement; positional information is encoded by the order of firing.

Alternatively, position could be represented by activation level, such as by an activation gradient, where the first letter is the most highly activated, the second letter is the next most highly activated, and so on, as in the SOLAR model (Davis, 1999). A bigram unit CR could then be tuned so that it responds when C's activation is higher than R's, but not vice versa. This is a logical possibility for how position-independent letter units could activate bigram units in parallel. This requires a monotonic gradient, and so the final letter would have the lowest activation. However, while letter perceptibility decreases across the string, it increases for the final letter (Rumelhart & McClelland, 1982; Lefton, Fisher & Kuhn, 1978). Where could this final-letter advantage originate, if the final letter is the least activated? Perhaps it is the case that identification accuracy is orthogonal to a letter's activation level, (i.e., accuracy only depends on whether or not the correct letter becomes activated, and is unrelated to subsequent activation level). Letter perceptibility would then be determined by the robustness of activity patterns at the sub-letter, featural level. However, as discussed in more detail below, perceptibility patterns for letter strings radically differ from those of non-letter symbols; accuracy for non-letter symbols depends directly on acuity, with no

²Such a dyslexia has not been encountered in more commonly studied languages, such as English. However, Hebrew orthography is particularly conducive to revealing a deficit of this sort, since vowels are not explicitly represented. Therefore, if the order of consonants is misperceived, there is a high probability that a word corresponding to the erroneous ordering exists for some combination of vowels. Thus, lexical constraints are reduced, allowing a pure deficit in position-encoding to be revealed.

advantage for initial and final symbols (Hammond & Green, 1982; Mason, 1982). Therefore, (standard) feature-level activation patterns cannot explain the letter perceptibility data at all.

In sum, a perceptibility function based on a letter-level activation gradient comes closer to explaining the data than feature-level influences, but is inconsistent with the final-letter advantage. While a letter-level activation gradient is not necessarily ruled out by the final-letter advantage, it is unclear what mechanism would give a perceptibility pattern that mirrors the activation gradient, except at the final letter. Davis (1999) suggests that under the SOLAR model, this pattern could be akin to the combination of a primacy gradient and a recency advantage seen in list processing. However, these phenomena arise under strictly serial presentation, and it is unclear how they would apply to SOLAR model, where all letters fire in parallel after the final letter is activated. We will see below how the particular serial mechanism of the SERIOL model, in which a *feature*-level activation gradient induces the serial firing, yields letter activations that decrease across the string, but rise for the final letter. Thus a serial encoding provides a way to bind letter-position information to position-independent letter detectors that allows congruence between letter activation levels and perceptibility patterns.

Furthermore, there must be a mechanism to activate these letter units. It is well-known that early visual representations are retinotopic. Activation of position-independent letter representations requires abstraction away from the encoding of retinal location. This functionality is also provided by a serial encoding, in which space (i.e., a retinotopic representation) is mapped into time, forming a location-invariant representation. In the following section, we will see in more detail how this accomplished in the SERIOL model.

As mentioned in the Introduction, an alternative approach to forming location-invariant representations is to assume that receptive-field sizes gradually increase (Mozer, 1991). At each successive processing level, locational information is reduced, while the complexity of detected features is increased. Could this approach be used instead to form a location-invariant, letter-level representation? No, if locational information were lost, there would be no information about letter position. For example, consider a “C” detector that

responds to the occurrence of a C across a wide range of retinal locations. Once this “C” detector is activated, information about the location/position of the C is lost, as there is no way to dynamically bind such information to the “C” detector in this scenario. Therefore, information about letter order could not be recovered.

This approach would instead require location-specific letter detectors - i.e. one “C” detector that responded to a small range of locations, another “C” detector for another range of locations, etc. Locational information would then be encoded by which “C” detector(s) responded. Then, there would have to be duplicated bigram detectors for all possible combinations of locations. For example, (using an abstract numbering of locations, increasing from left to right), a CR detector that responds to C in location 1 and R in location 2, 3, or 4, another CR detector for C in location 2 and R in location 3, 4, or 5, etc. It would not be possible to use a single bigram detector for different combinations because there would be no way to ensure that the C occurred to the left of the R. Thus such an approach is inconsistent with the evidence for position-independent letter units, and is inefficient. Furthermore, it is unclear how this scheme could handle different font sizes, which would affect the locations of the constituent letters. In contrast, a serial encoding requires only a single detector for each bigram, which fires whenever the constituent position-independent letter detectors are activated in the correct order. Such a scheme is not sensitive to the size of the letter stimuli, as locational information has already been abstracted away.

Note also that the proposal that “receptive field sizes gradually increase” can be used in a purely descriptive way. It could be said that receptive-field size increases in going from the letter level to the bigram level in the SERIOL model, as bigram nodes represent a larger area of space than do letter nodes. However, this tells you nothing about how bigrams become activated. In sum, if you have position-independent letter detectors, there must be some way to bind locational/positional information to such units. Any such mechanism would go beyond the proposal that location-invariance is directly achieved simply by increasing receptive field sizes. I also note that, in an implemented model of the idea that increases in receptive-field

size directly allow location-invariance (Mozer, 1991), the problem of how to bind position information to letter representations was finessed by using a large jump in receptive field size, from features directly to multi-letter units. Thus, there was no level of representation corresponding to letters, which is inconsistent with the evidence for position-independent letter detectors.

Interaction with Phonology

Thus far, I have focused on the direct, orthographic route to the lexicon. However, the encoding of letter position also supports an assembled phonological encoding of the string. This, of course, depends on having learned the patterns of grapheme-phoneme correspondences. Simulations have indicated that the method of letter-position encoding affects the learnability of such relationships. A simulation based on parallel, multi-letter units (trigrams) showed poor generalization on reading pseudowords (Seidenberg & McClelland, 1989). This was likely due the choice of encoding; representations of the same letter in different contexts bore no relationship to each other, making generalization difficult. This is known as the *dispersion problem*. Changing the input encoding to a slot-based one which encoded graphemes in onset, vowel and coda positions allowed better performance on pseudoword reading (Seidenberg, Plaut, Petersen, McClelland & McRae, 1994).

However, such a slot-based encoding is not consistent with human performance. That is, if a slot encoding does not explicitly encode the order of letters within the onset and coda, it cannot account for the fact that we can detect misorderings, such as SRTING. If a slot encoding does explicitly encode order via different slots for different positions (e.g. coda-1 = S, coda-2 = T, coda-3 = R), this is an absolute-position coding and is inconsistent with the relative-position and transposition priming data (Humphreys et al., 1990; Grainger & Whitney, 2004; Peressotti & Grainger, 1999; Perea & Lupker, 2003). (Note that this does not rule out an onset-coda representation at a higher phonological level; rather the argument is against such a scheme as the basic encoding of letter order).

Parallel activation of bigram units would also lead to the dispersion prob-

lem. Thus, such a representation is not suitable for learning phonological relationships. In contrast, a serial encoding of letter order is highly suitable for this task. It is based on position-independent letter units, so there is no dispersion problem. In addition, phonology is inherently serial. Temporal alignment of a serial orthographic and phonological representations could facilitate the learning of the relationships between the two, especially during early reading acquisition, where simple words having one-to-one letter-to-phoneme correspondences dominate. I have proposed that the failure to form a serial representation of letter order could be a contributing factor to phonological problems observed in development dyslexia (Whitney & Cornelissen, 2005). This proposal is consistent with MEG evidence for early abnormal visual processing of letter strings (at about 150 ms post-stimulus) in dyslexics (Helenius, Tarkiainen, Cornelissen, Hansen & Salmelin, 1999). Furthermore, others have presented evidence that phonology is assembled serially during visual word recognition, implying serial activation of letter units (Perry & Ziegler, 2002; Roberts, Rastle, Coltheart & Besner, 2003). As such a representation encodes letter order, it would be redundant for the orthographic route to use a different encoding.

Thus I assume that processing splits after the letter level. Along the phonological route, a serial letter encoding activates a serial phonological representation, which then activates the lexical level, perhaps using the same mechanisms as for recognizing speech. Along the orthographic route, the serial letter encoding activates open-bigram units, which also contact the lexical level.

Summary

The abandonment of the serial assumption seems to have been based on unwarranted interpretations of lexical-decision data (Frederiksen & Kroll, 1976; Richardson, 1976) and the WSE (Reicher, 1969), and on a general bias toward parallel processing, which was largely driven by the popularity of a certain style of modeling (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1986). Thus, there is no evidence *against* serial lexical access.

In contrast, there is direct evidence for serial read out of letter strings (Nice & Harcum, 1976). Additionally, a serial encoding provides binding and abstraction mechanisms, allowing the formation of an abstract location-invariant representation of letter order based on position-independent letter units. Such a representation is suitable for learning grapheme-phoneme relationships due to its dispersion-free nature, and its congruence with the serial nature of phonology. Thus, there are good arguments and compelling evidence *for* a serial encoding.

Review of SERIOL Model

Having discussed the motivations for a serial encoding, I now review the SERIOL model (Whitney, 2001; Whitney & Berndt, 1999) in more detail. This will then allow demonstrations of how the serial mechanisms account for specific patterns of experimental data.

The SERIOL framework (Sequential Encoding Regulated by Inputs to Oscillations within Letter units) is a comprehensive account of string processing in the proficient reader. It offers a computational theory of how a retinotopic representation is converted into an abstract representation of letter order. The model consists of five layers: edge, feature, letter, bigram, and word. We have already seen that there is a serial encoding at the letter level, which activates the bigram level. But how is this serial encoding induced? I will give a brief overview of the underlying principles, and then discuss each layer in more detail.

First, some terminology. Each layer is comprised of processing units called *nodes*, which represent groups of neurons. Within each layer, the activation of a letter is taken to be the total amount of activity occurring across all nodes representing that letter. Thus, activation depends on the number of such nodes, their firing rate, and their firing duration.

The serial activation of letter nodes is induced by an activation gradient from the feature level, dubbed the *locational gradient*, which decreases across the string (i.e., the first letter’s features are the most highly activated, the second letter’s features are the next most, etc.) Additionally, it is assumed

that letter nodes undergo synchronous sub-threshold oscillations of excitability. The interaction of the changing excitability of the letter nodes, and the varying levels of bottom-up input leads to serial firing of the letter nodes.

At the lowest layer of the model (the edge layer), there is a different activation pattern, resulting from visual acuity. This acuity gradient is congruent with the locational gradient in one cerebral hemisphere, but not the other. Therefore, hemisphere-specific processing is required to create the locational gradient. Next, I discuss this processing in more detail.

Edge to Feature Layers

The lowest layer of the model, the edge layer, corresponds to the early visual cortical areas. Properties of the edge level reflect the known architecture of the visual system. The edge layer is retinotopically organized, and is split along the vertical meridian, corresponding to the two cerebral hemispheres. The number of edge nodes representing a letter is taken to decrease as eccentricity increases, corresponding to the acuity gradient. Therefore, letter activations at the edge level decrease as eccentricity increases.

The feature layer of the model is also retinotopically organized and split across the hemispheres. The activation pattern from the acuity gradient must be converted to the locational gradient as the edge layer activates the feature layer. Note that for a fixated word, the acuity pattern across the letters in the RVF/LH is the same as the locational gradient (i.e., decreasing from left to right). Therefore, the acuity gradient can serve as the locational gradient for those letters. However, in the LVF/RH, the acuity gradient *increases* from left to right. Therefore, when the edge level activates the feature level, the acuity gradient must be inverted in the LVF/RH. Next we consider the details of this processing, which is assumed to be learned during reading acquisition in response to a top-down attentional gradient (Whitney & Cornelissen, 2005).

The locational gradient is created via modification of firing rates, as displayed in Figure 1. I propose that letter features in the LVF/RH become more highly activated by edge-layer inputs than those in the RVF/LH. This

allows the first letter to reach a high level of activation. Within the RH feature layer, I propose that there is strong left-to-right lateral inhibition. That is, a feature node inhibits nodes to its right. As a result, features comprising the first letter receive no lateral inhibition, and inhibition increases as letter position increases. Thus, the first letter attains the highest activation level (due to strong excitation and lack of lateral inhibition), and activation decreases toward fixation (due to sharply increasing lateral inhibition as the number of letters on the left increases).

In the RVF/LH, the acuity gradient can serve as the locational gradient. Overall excitation is weaker than to the LVF/RH. Left-to-right inhibition is not necessary, although some weak such inhibition may steepen the slope of the gradient.

The two hemispheric gradients are “spliced” together via cross-hemispheric inhibition. The RH features inhibit the LH features, bringing the activation of the LH features lower than the activation of the least activated RH features. As a result, an activation gradient that is strictly decreasing from left to right is created.

Feature to Letter Layers

Next we consider in more detail how the locational gradient induces serial firing at the letter layer. Hopfield (1995) and Lisman and Idiart (1995) have proposed related mechanisms for precisely controlling timing of firing, in which nodes undergo synchronous, sub-threshold oscillations of excitability. The amount of bottom-up input then determines timing of firing with respect to this oscillatory cycle. The interaction of the feature and letter layers of the model are based on this proposal. The letter nodes comprise the oscillatory nodes, which are taken to have a cycle length of about 200 ms (i.e., in the theta range). It is assumed that the oscillatory phase is time-locked with the activation of the feature level, such that the feature level becomes active during the trough of the oscillatory cycle. Such an assumption is consistent with MEG evidence that theta phase is reset by anticipation of stimulus presentation (Tesche & Karhu, 2000), indicating that there are

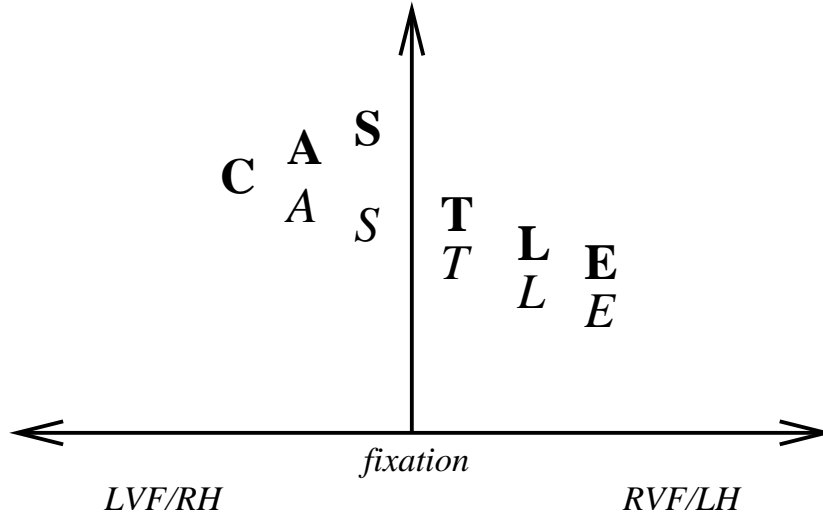


Figure 1: Formation of the locational gradient at the feature layer, for the centrally fixated stimulus CASTLE. The horizontal axis represents retinal location, while the vertical axis represents activation level. The bold-face letters represent bottom-up input levels, which are higher in the RH than the LH. In each hemisphere, activation decreases as a eccentricity increases, due to the acuity gradient. The italicized letters represent the effect left-to-right inhibition within the RH, and RH-to-LH inhibition in the LH. In the RH, C inhibits A, and C and A inhibition S, creating a decreasing gradient. The RH inhibits each letter in the LH by the same amount, bringing the activation of T lower than that of S. As a result, activation monotonically decreases from left to right.

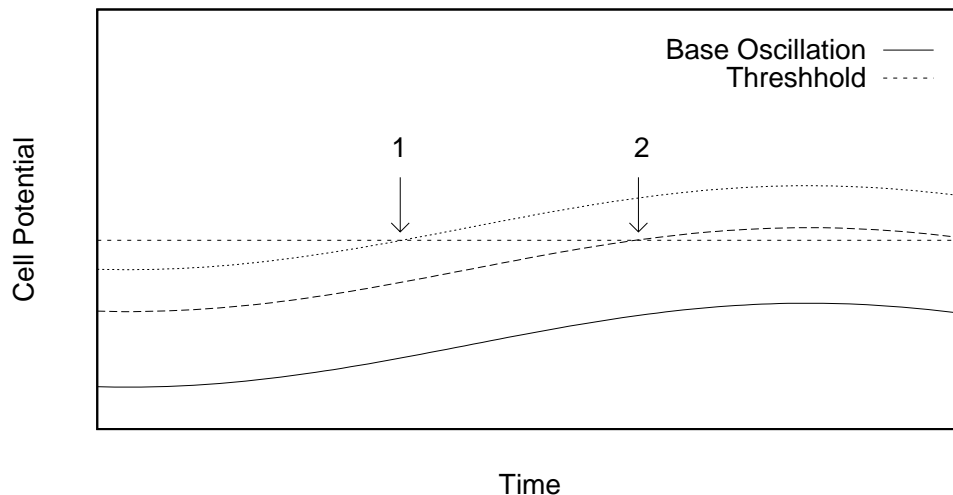


Figure 2: Interaction of input level and timing of firing for a cell with undergoing a sub-threshold oscillation of excitability. When a relatively high level of input (top curving line) is added to the base oscillation, the cell crosses threshold at time 1 (action potential not illustrated). If less input were received, the cell would cross threshold later in the cycle, such as at time 2.

brain mechanisms to coordinate input and oscillatory phase. (Under normal reading conditions, theta phase may be coordinated with saccade generation, for example.) However, the mechanisms underlying this time-locking are beyond the scope of the model.

Input level then determines how early in the cycle such a letter node can cross threshold and fire. (See Figure 2.) Near the beginning (trough) of the cycle, excitability is low, so only a node receiving a high level of input can cross threshold and fire. Excitability increases over time, allowing nodes receiving less and less input to progressively fire. In our CART example, the C node would receive the most input, A the next, R the next, and T the least, allowing C to fire the earliest, A next, R next, and finally T.

An activated letter node inhibits other letter nodes. As a letter node continues to fire, its firing rate slows, reducing lateral inhibition to the other nodes. This allows a new letter node to start firing. When an active letter node receives lateral inhibition, it then becomes strongly inhibited, so that

it will not re-fire for the remainder of the oscillatory cycle.

This process also creates varying activation levels. The activation of a letter node depends on both its rate and duration of firing. Under the assumptions that a higher input level leads to faster firing, and that firing duration is fairly constant across letters, the result is a decreasing activation gradient at the letter level. However, the node representing the final letter is not inhibited by a subsequent letter. It can continue fire until the end (down-phase) of the oscillatory cycle. Therefore, the final letter can fire longer than the other letters, and reach a higher level of activation than the internal letters even though it receives less input. This explains the final-letter advantage. As discussed below, this proposal also explains the counterintuitive experimental results on letter perceptibility.

Letter to Bigram Layers

A bigram node XY becomes activated when letter node X fires, and then letter node Y fires within a certain time period. A bigram node responds with a burst of firing, and then is quiet. The number of spikes in this burst decreases as the time increases between the firing of X and Y. That is, the activation of bigram XY is highest when triggered by contiguous letters, and decreases as the number of intervening letters increases. Priming data indicates that the maximum separation is likely to be two letters (Schoonbaert & Grainger, 2004).

In previous articles on the SERIOL model, I have assumed that bigram activation levels were also influenced by letter activation levels, which are influenced by string position (Whitney, 2001; Whitney & Berndt, 1999). However, this overall assumption is inconsistent with emerging evidence on weak positional effects of priming at the word level (Granier et al., 2005), as discussed below. Therefore, I now take bigram activation levels to be affected only by the separation of the constituent letters.

Following the evidence for a special role for external letters (Humphreys et al., 1990), the string is anchored to these endpoints via edge bigrams.³

³This is a new assumption. The importance of the external letters was formerly cap-

That is, bigram $*X$ is activated when letter X is preceded by a space, and bigram Y^* is activated when letter Y is followed by a space. In contrast to other bigrams, an edge bigram cannot become partially activated (e.g., by the second or next-to-last letter). I assume a special mechanism for the activation of edge bigrams, which operates somewhat differently than for bigrams detecting a pair of letters. The details of this edge detection are left for future work.

Because letters are activated sequentially, bigram activations occur sequentially. For example, the input *cart* first activates bigram node $*C$ (when letter node C fires), then CA (when A fires), then AR and CR (when R fires), then RT , AT , and CT (when T fires), and then T^* .

Bigram to Word Layers

Bigram nodes connect to word nodes via weighted connections. The weight on a bigram-word connection is proportional to the activation level of that bigram when that word is presented as input (as would result from Hebbian learning.) As is usual in neural network models, the weight vector is normalized, so that bigrams making up shorter words have higher connection weights than bigrams making up longer words. For example, this allows the string *tee* to activate the word node TEE more than the word node $TEETHE$.⁴

The input to a word node is the dot product of the weight vector and input vector. The input vector changes over time, because bigram activations occur serially, as indicated above. The activation of a word node at time t is a function of its activation at time $t - 1$ and the input at time t . Lateral inhibition within the word layer also operates over time. That is, as the bigrams fire, there is ongoing competition between word nodes.

tured via high activations of bigrams containing those letters. However, now that bigram activation levels do not reflect letter activation levels, edge bigrams are now assumed.

⁴Normalization is another new assumption. Information concerning the length of the string was formerly carried on the activation levels of bigrams containing the final letter.

Summary

As edge nodes activate feature nodes, the acuity gradient is transformed into the locational gradient, via hemisphere-specific processing. The locational gradient interacts with oscillatory letter nodes to induce serial firing of letter nodes. A bigram node fires when the constituent letter nodes fire in the correct order. Bigram nodes contact the word level via weighted connections. As word nodes are progressively activated by the bigram nodes, there is ongoing lateral inhibition between word nodes.

While the model may seem somewhat complex, keep in mind that the goal of the model is to specify how the early retinotopic representation of a string could be progressively transformed into an abstract encoding of letter order, and to do so in way that is neurobiologically plausible and that is consistent with experimental data. The central serial mechanism is based on computational models proposed by leading neurobiologists (Hopfield, 1995; Lisman & Idiart, 1995), and is consistent with a wide range of evidence for the importance of oscillatory activity in the brain. It should not be surprising that such a comprehensive, realistic model of brain function route would yield some complexity.

Letter Perceptibility Evidence for SERIOL Model

Having reviewed the model, I now discuss some experimental support for the particulars of the model based on novel analyses of previous studies of unilateral perceptibility patterns. The following explanations of these patterns go beyond the general assumption of seriality; they depend on proposals specific to the SERIOL model for *how* the serial encoding is created.

Note that the model was not explicitly designed to explain the unilateral patterns. Rather, the mechanisms in the model were chosen to solve the representational problems of converting a retinotopic representation of a fixated string into an encoding of relative letter position. These mechanisms also explain the unilateral perceptibility patterns. While the following explanations may seem to require a lot of machinery, keep in mind that these mechanisms

are specified by a pre-existing model.

Letter Perceptibility Patterns

It is well-known that under central fixation, the external letters of a string (first and last letters) are the best perceived (Rumelhart & McClelland, 1982). The standard explanation of this pattern is that the external letters undergo less lateral masking than the internal letters and therefore have a perceptual advantage. However, this explanation is not consistent with the perceptibility pattern for non-letter strings. Centrally fixated strings of five non-letter symbols do not yield an exterior-symbol advantage. Rather, perceptibility monotonically decreases with eccentricity. The first and last symbols are the least well perceived, as would be expected on the basis of acuity, whereas letter stimuli yield the usual exterior-letter advantage under the same experimental conditions (Hammond & Green, 1982; Mason, 1982). Note that these experiments were conducted using a between-subjects design for stimulus type. That is, one group of subjects saw letters, while another group saw symbols. Therefore, mechanisms specific to letter strings would only be activated in those subjects seeing letters, so the problem of how the feature level could “know” whether to invoke string-specific processing on a trial by trial basis is not an issue.

If the external letter advantage actually arose from reduced low-level lateral masking, such an advantage should also be present for non-letter symbols, but it is not. Rather, the external letter advantage arises from processing that is specific to letter strings.

For unilateral presentation of short strings (3 or 4 letters) at large eccentricities, a counterintuitive pattern arises. In the LVF/RH, the first letter is the best perceived of all the letters; in the RVF/LH, the last letter is the best perceived (Bouma, 1973; Estes, Allemeyer & Reder, 1976; Legge, Mansfield & Chung, 2001). Thus, in each visual field, the letter *farthest* from fixation (where acuity is the lowest) is the most likely to be correctly reported. This pattern is present even at long exposure durations. For example, see Figure 3.

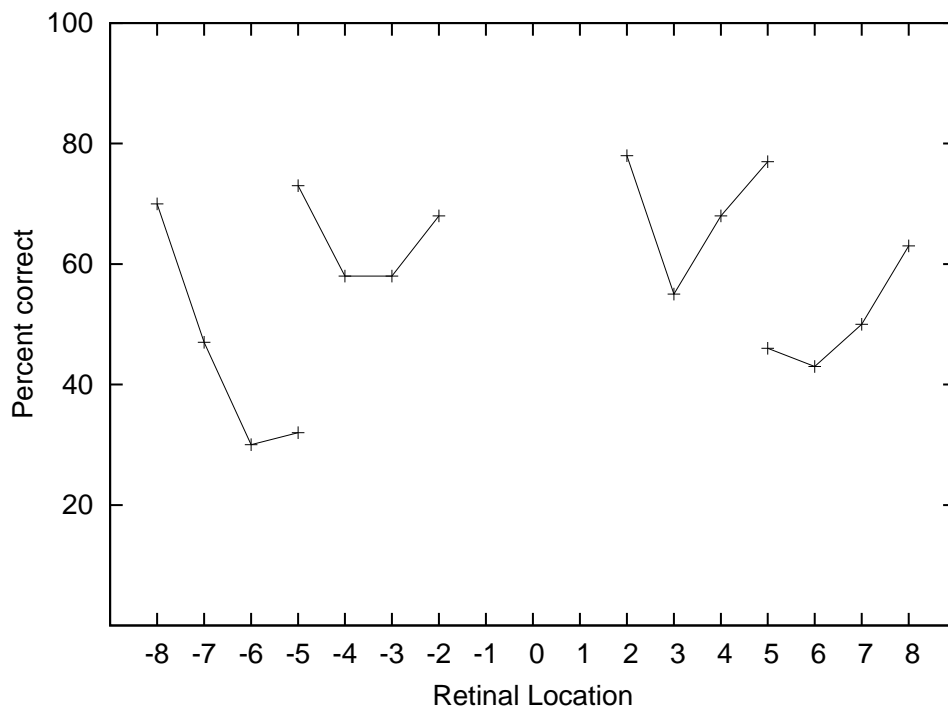


Figure 3: Experimental results regraphed from Estes and colleagues (1976) Figure 3 for a four-letter string, occurring across two different retinal locations (given in letter widths) in each visual field. Exposure duration was 2400 ms. (Subjects were trained to maintain central fixation, and their gaze was monitored.) Results are for the no-change condition, where a mask of filler characters (all \$'s or all #'s) across retinal locations -9 to 9 was initially displayed, and then four of those characters were changed to letters, while the remainder of the filler characters remained the same.

For non-letter symbols, it has been shown that lateral masking is somewhat stronger from the peripheral side than from the foveal side (Chastain, 1989). That is, for a target and a single distractor, accuracy is worse when the distractor occurs to the side away from fixation than when it occurs to the side nearer fixation. Perhaps, under lateralized presentation, perceptibility is poor for the inner external letters (last letter in LVF, first letter in RVF) because they are peripherally masked, while the outer external letters (first letter in LVF, last letter in RVF) are not? This explanation is unlikely. Note that the lateral-masking results are for a target at a fixed eccentricity, where the location of the distractor is varied. In contrast, the inner letters are considerably closer to fixation than the outer letters. As discussed above, initial and final non-letter symbols are perceived less well than internal symbols closer to fixation (Hammond & Green, 1982; Mason, 1982), despite the fact that an external symbol is not masked peripherally, while an internal symbol is masked both peripherally and foveally. This result indicates that the effect of acuity is stronger than any effect of lateral masking, and so peripheral masking cannot explain why the letters closest to fixation are poorly perceived. Furthermore, as discussed in more detail below, the effect of changing the surrounding characters in the Estes et al. (1976) experiments (versus not changing them) was strongest for LVF initial letters. This specificity is predicted by the SERIOL model, and cannot be explained by a general effect of lateral masking.

Rather, I propose that these perceptibility patterns are a direct result of visual processing that is unique to letter (and number) strings. As discussed above, the induction of the serial encoding leads to differing activation levels across the letter layer. In general, I assume that perceptibility indexes a letter's activation level. This is a simplification, as it essentially assumes that pre-conscious recognition of a letter always proceeds correctly, and activation level then determines the probability of availability for report. For example, it assumes that when an 'u' is presented, the letter node 'U' is correctly activated, and any inability to report 'U' arises from a low activation level of that letter node. Actually, activation level should be weighted by the probability of correct pre-conscious recognition, which would directly depend on acuity.

However, I assume that for nearby letters, differences in activation levels have a larger effect than differences in acuity. This simplifying assumption is consistent with the fact that letter-perceptibility patterns radically differ from what would be expected on the basis of acuity (Hammond & Green, 1982; Mason, 1982), in contrast to non-alphanumeric symbols.⁵

Recall that letter-layer activation level is function of firing rate and firing duration, which are determined by bottom-up input levels to the target node and to the next letter node, respectively. These input levels arise from the locational gradient. As discussed next, the shape of the locational gradient varies with presentation location. These differing shapes of the locational gradient and the ensuing letter-layer firing patterns then explain the above perceptibility patterns.

Locational Gradient Shapes

Recall that locational-gradient formation requires differential processing across the hemispheres. In the RVF/LH, the acuity gradient serves as the locational gradient. In the LVF/RH, the acuity gradient is inverted via strong bottom-up excitation and left-to-right lateral inhibition. Because the locational gradient is formed by different mechanisms in each hemisphere, the shape of the resulting gradients may vary across presentation locations, especially when large eccentricities are considered.

The acuity gradient is known to fall off less quickly as distance from fixation increases (Westheimer, 1987). That is, the slope of the acuity gradient is steepest near fixation, and becomes shallower as eccentricity increases.⁶

⁵However, recognition probability may well have a sizable effect near fixation. As discussed in more detail in the following section, the difference in acuity of contiguous letters is largest when one of the letters falls directly at fixation. In this case, the effect of the much higher recognition probability for the fixated letter may dominate. This explains why a fixated letter is perceived better than the letter just to the left (denoted the *leftward letter*) (Mason, 1982; Wolford & Hollingsworth, 1974). That is, within each trial, the activation level of the letter node activated by the leftward letter may actually be higher than that of the fixated letter, but the probability across trials that the correct letter node is activated by the leftward letter is significantly lower than for the fixated letter.

⁶This acuity pattern is commonly misrepresented as “acuity falls off rapidly outside the fovea”, implying that acuity is uniformly high across the fovea and then falls off. This is not

Because the RVF/LH locational gradient is based on the acuity gradient, this implies that the RVF/LH locational gradient becomes more shallow as eccentricity increases. (See right half of Figure 4.)

In the LVF/RH, formation of the locational gradient depends on left-to-right lateral inhibition. This processing is optimized to create the locational gradient for a small number of letters near fixation. For longer strings at large eccentricities, inhibition may be too strong at early string positions (due to their relatively low level of activation), but may become too weak at later string positions (due to the increasing acuity). (See left half of Figure 4). Thus the prediction is that the locational gradient should vary with visual field. Indeed, this account explains observed differences in perceptibility patterns (Wolford & Hollingsworth, 1974), as demonstrated by a computational model (Whitney, 2001). Moreover, this theory of locational gradient shape has lead to experimentally verified predictions as to how to reverse VF asymmetries associated with string length and orthographic-neighborhood size (Whitney & Lavidor, 2004, 2005).

Note that the proposed stronger bottom-up excitation for the LVF/RH is not inconsistent with the well-known RVF advantage observed for lexical tasks. The LVF/RH left-to-right inhibition outweighs this excitation for non-initial positions, so that input levels are lower for non-initial LVF letters than RVF letters. Furthermore, the resulting LVF/RH activation pattern is not smoothly decreasing, providing a non-optimal encoding of letter order (especially for longer words) and decreased activation of the target word at the lexical level.

Account of Perceptibility Patterns

For a centrally fixated string, the initial-letter advantage and final-letter advantage arise for different reasons in the SERIOL model. The initial-letter node has an advantage because it receives the highest level of bottom-up input, allowing it to fire the fastest. It receives the most input because, at the

the case. Rather, acuity falls off rapidly within the fovea, so that acuity is substantially reduced by the fovea / parafovea boundary. The rate of decrease in acuity is actually sharper across the fovea than the parafovea.

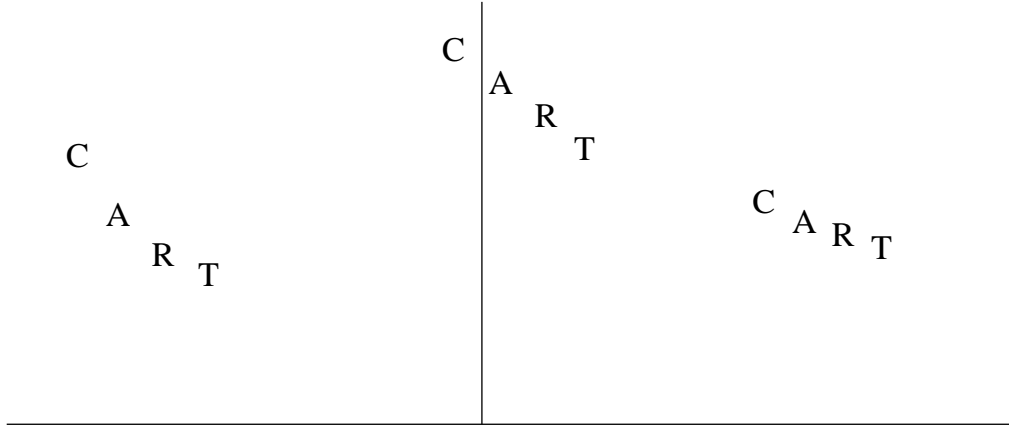


Figure 4: Schematic of locational gradients for the stimulus CART at three different presentation locations. The vertical axis represents activation, while the horizontal axis represents retinal location. For central presentation, the gradient is smoothly and rapidly decreasing. For RVF presentation, the gradient is shallower because the acuity gradient is shallower. For LVF presentation, the initial letter strongly inhibits nearby letters, but the gradient flattens out as acuity increases.

feature level, it is not inhibited from the left. The final letter has an advantage because it is not inhibited by a subsequent letter during the induction of serial firing. That is, it is not inhibited from the right at the letter level. Thus, like others, I also attribute the advantage for the external letters to a lack of lateral inhibition. However, this reduced inhibition does not arise from less masking at a very low level (as is generally assumed). Rather it arises from mechanisms specific to the conversion of a spatial array of letters into a serial encoding of order.

The present analysis of the external-letter advantage implies that it should be possible to differentially affect the initial- and final-letter advantages. The initial-letter advantage should disappear if the amount of bottom-up input to the first letter node is not significantly higher than to the other letters. The final-letter advantage should disappear if the last letter node starts firing late in the oscillatory cycle, and so is unable to fire for a longer time than the other letters.

As we see next, these proposals explain the counterintuitive perceptibil-

ity patterns for lateralized presentation of short strings. In particular, we will consider the results of Estes and colleagues (1976), as given in Figure 3. As noted in the figure’s legend, the letter stimuli replaced some characters in a string of masking characters. Thus, when the letter stimuli appeared, only those locations containing the letters changed identity. Visual attention would then be automatically focused on the locations where changes occurred, minimizing the effect of the surrounding static mask characters. Therefore, in the following analysis, I assume that there is little effect of the mask characters at the feature level. As discussed below, a different perceptibility pattern arose in other trials where the mask characters also changed identity when the letter stimuli appeared.

In the following discussion, *primacy* will signify that a letter is perceived better than all other letters, whereas *advantage* will mean that an external letter is perceived better than the internal letters. First we consider LVF presentation. Within the feature layer, the initial letter is strongly excited, and strongly inhibits letters to the left, causing them to have low activation levels. At the letter level, the first letter can fire quickly (due to a relatively high level of bottom-up input) and for a relatively long time (because it is not rapidly cut off by the next letter, as the next letter receives a much lower level of bottom-up input). So the initial letter attains a high activation level. Due to the low input level to the final-letter node, it starts firing late in the oscillatory cycle. Therefore, it cannot fire longer than the other letters, and no final-letter advantage emerges. This explains the perceptibility pattern for locations -8 to -5. (See Figure 5.)

For RVF presentation, overall bottom-up excitation to the feature layer is weaker. Therefore, the activation of the initial letter’s features is not boosted to a high level. Furthermore, there is weak left-to-right inhibition, while the acuity/location gradient is quite shallow, so the activation of the second letter’s features is quite close to that of the first letter. As a result, at the letter level, the firing of the first letter is rapidly cut off by the second letter, giving no initial-letter advantage. Each successive letter quickly inhibits the preceding letter, allowing the final letter to start firing early in the oscillatory cycle. Therefore the final letter can fire longer than the other letters, creating

a final-letter primacy. This explains the perceptibility patterns for locations 5 to 8. (See Figure 5.)

This account also explains the initial/final difference within a single retinal location (at -5 and 5 in Figure 3). In the LVF/RH, the left-to-right, feature-level inhibition creates a disadvantage for a final letter, whereas an initial letter does not receive this inhibition. In the RVF/LH, the shallow locational gradient creates a disadvantage for an initial letter because its firing at the letter layer is rapidly inhibited by the second letter. For a final letter, firing at the letter layer can instead continue until the end of the oscillatory cycle.

In contrast to the larger eccentricity, the perceptibility function is U-shaped for -5 to -2 and 2 to 5. Due to higher acuity, bottom-up input is higher overall. In the LVF/RH, this allows the final letter to start firing earlier in the cycle, creating a final-letter advantage. Along with the usual initial-letter advantage, this gives the U-shaped pattern. In the RVF/LH, the acuity/locational gradient is steeper than for the larger eccentricity, so the difference in input to the first and second letters is larger, creating an initial-letter advantage and giving an overall U-shape.

As indicated above, there were also trials in the experiment reported by Estes and colleagues (1976) in which the masking characters changed identity when the letter stimuli appeared (change condition). In this case, changes occurred at all retinal locations, and visual attention could not automatically be directed to the letters. Therefore, the presence of the surrounding mask characters would have a stronger effect at the feature level. The SERIOL model predicts, however, that this effect should vary with visual field and string position. Recall that the initial-letter primacy in the LVF specifically depends on a lack of left-to-right inhibition at the feature level. Therefore, increased interference at the feature level should preferentially degrade perceptibility of initial letters in the LVF. That is, external letters would be expected to suffer some general decrement in perceptibility due to increased non-specific lateral masking, but an LVF initial letter should suffer an additional disadvantage because it is no longer uninhibited by features to the left.

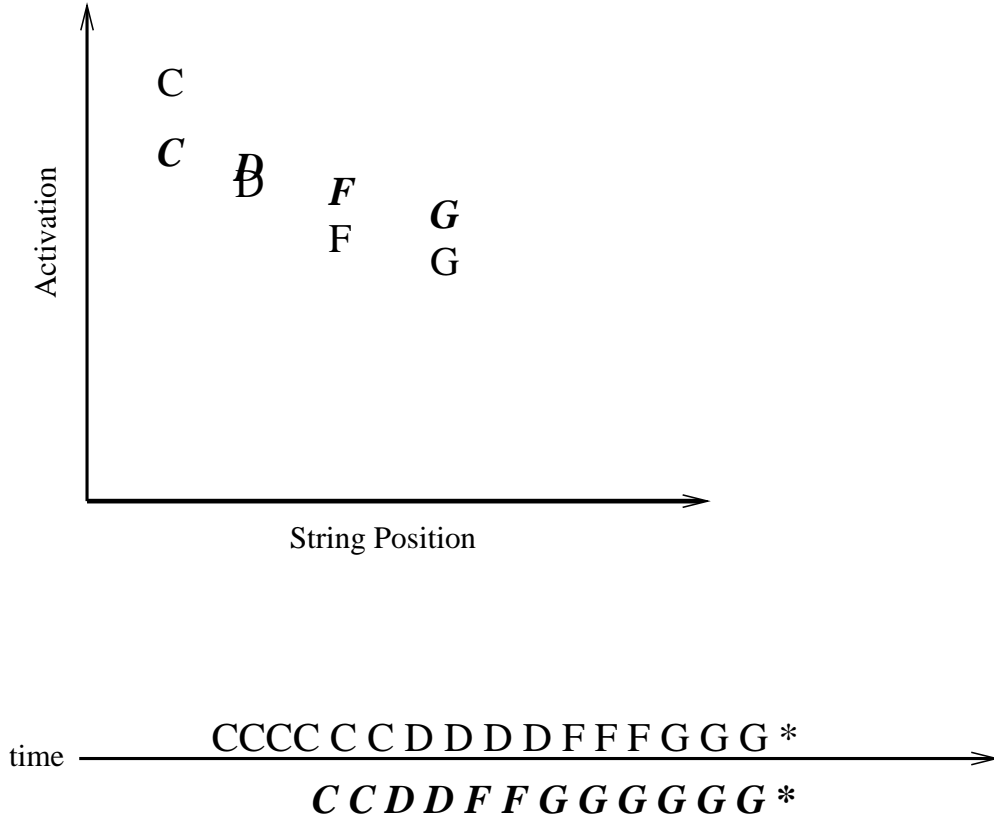


Figure 5: Locational gradients and resulting firing patterns for LVF/RH (normal font) and *RVF/LH* (bold italics) presentation. Top: Comparison of locational gradient for string CDFG under RVF/LH presentation and LVF/RH presentation. Bottom: Cartoon of resulting firing pattern at the letter level. The point in the oscillatory cycle at which the down phase prevents further firing is marked *. In the LVF/RH, the first letter fires faster and longer than the other letters, because it receives a much higher level of input. The variations in the amount of bottom-up input creates decreasing activation across the string. The final letter starts firing late in the cycle, and is soon cut off by the end of the oscillatory cycle, giving no final-letter advantage. In the RVF/LH, each letter rapidly cuts off the previous letter, allowing the final letter to fire a long time. As a result, activation is flat across the string and rises for the final letter. These firing patterns account for the perceptibility patterns at the larger eccentricities in Figure 3

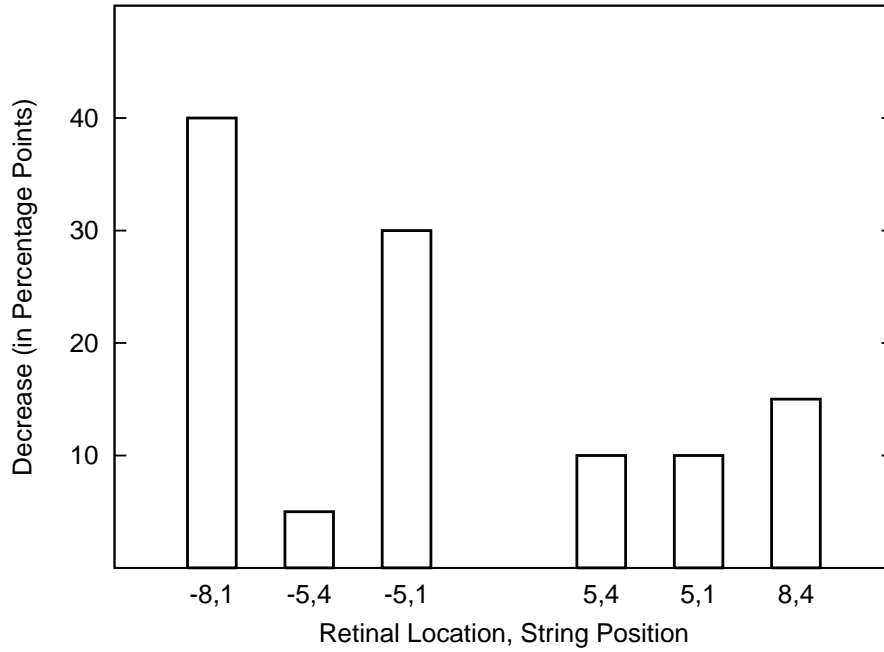


Figure 6: Effect of the change condition relative to the no-change condition for external letters under 2400 ms exposure duration (i.e., percent correct in no-change condition minus percent correct in change condition). Values were read off of Figure 3 (Estes et al., 1976), rounding to nearest 5 percentage points. At retinal locations 2 and -2 the change condition had no effect, likely due to the high acuity.

This is exactly what occurred. Figure 6 shows the effect of the change condition relative to the no-change condition. Perceptibility for LVF initial positions decreased by 30 percentage points or more, while perceptibility for all other location/position combinations decreased by 15 percentage points or less. In particular, the effect of the change condition was much stronger for an initial letter than for a final letter at -5 (LVF), but did not vary with string position at 5 (RVF). This interaction was statistically significant (Estes et al., 1976).

The effect of string position at -5 could not be due to lateral masking that is stronger in the peripheral direction than the foveal direction, because this effect was not present at 5. Rather, this pattern is predicted and explained by the SERIOL model's proposal that there is strong left-to-right, feature-level

inhibition in the LVF/RH that is specific to string processing, and that the LVF initial-letter primacy arises directly from the lack of such inhibition.

Dependence on Exposure Duration

We have seen how interactions at the feature level preferentially affect initial letters in the LVF. Next we now consider how final letters in the RVF could be preferentially affected. The RVF final-letter primacy depends on the ability of the final letter to start firing relatively early in the oscillatory cycle, so that it can fire longer than the other letters. Therefore, if it were possible to shift the firing of all letters later into the cycle, the final letter should enjoy less of an advantage because it cannot fire for as long a time. An overall decrease in bottom-up activation levels would have such a shifting effect, and such a decrease could likely be brought about by decreasing exposure duration. Thus, we would *not* expect to see a RVF final-letter primacy at very brief exposures, because the overall low level of input pushes the firing of the final letter late into the oscillatory cycle. As exposure duration increases, the firing of all the letters is shifted earlier and earlier into the cycle, allowing the final letter to fire longer and longer. In contrast, the activation of a *non-final* RVF letter should not change much as exposure duration increases, because the firing of a non-final letter is still quickly cut off by the subsequent letter (due to the shallow acuity/location gradient). Thus, in the RVF, a final-letter primacy should emerge as exposure duration increases.

However, in the LVF, the initial-letter primacy *should* be present at very brief exposures, because it is based on strong left-to-right inhibition at the feature level, which does not depend on temporality. As exposure duration increases, the initial letter should be the primary beneficiary because, at the feature level, the increased bottom-up input to non-initial letters is canceled by increased lateral inhibition from the initial letter.

To summarize the predictions, in the RVF, the final-letter primacy should *not* be present at very brief exposures. Increasing exposure duration should primarily benefit the final letter, creating a final-letter primacy. In the LVF, the initial-letter primacy *should* be present at very brief exposures. Increasing

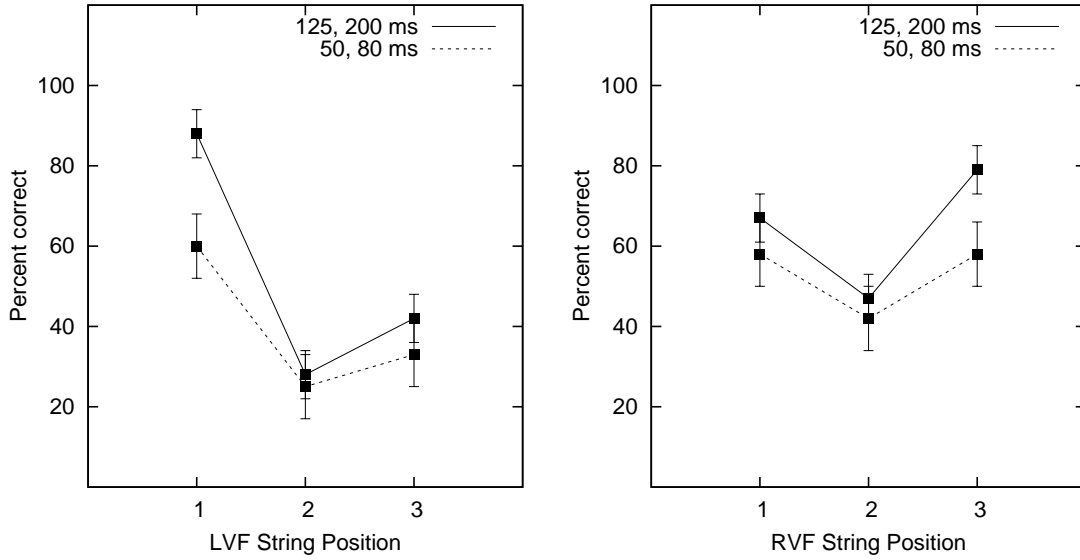


Figure 7: Results from Experiment 2 (Legge et al., 2001) for the two largest eccentricities, grouped by exposure duration, with 95% confidence intervals.

exposure duration should primarily benefit the initial letter, increasing its primacy.

A literature search revealed that a relevant experiment had already been performed, in which retinal location and exposure duration were systematically varied in a trigram identification task (Legge et al., 2001). However, the published data were not presented in a way that would allow evaluation of the above predictions, so I inquired about obtaining the raw data from the authors, who kindly provided it. The data were analyzed for the two largest eccentricities (-12 to -10 and -11 to -9 versus 9 to 11 and 10 to 12, in units of letter widths) for brief exposures (50 ms and 80 ms) versus longer exposures (125 ms and 200 ms). This analysis did indeed reveal the predicted patterns, as shown in Figure 7.

In summary, this account of perceptibility patterns relies directly on the seriality at the letter level. The lack of an initial-letter advantage for RVF presentation arises because the second letter quickly cuts off firing of the first letter, whereas the initial-letter primacy for LVF presentation is in part

due to the second letter not being able to quickly cut off firing of the first letter. The final-letter primacy for RVF presentation arises because the firing of the final letter is not cut off by a subsequent letter. The lack of a final-letter advantage for LVF presentation arises because the final letter starts firing late in the oscillatory cycle, and is quickly cut off by the down-phase of the cycle. As discussed above, these firing patterns at the letter level directly depend on specific proposals for *how* the serial encoding is created, based on hemisphere-specific processing in the formation of the locational gradient, and on the interaction of the locational gradient and the oscillatory cycle. The only premise that is required in addition to the basic principles of the model is that the acuity gradient is shallower in the parafovea than the fovea, which is a known property of the visual system. No other model can explain these counterintuitive perceptibility patterns, the dependence of the LVF initial-letter primacy on a lack of feature-level interference, and the dependence of the RVF final-letter primacy on exposure duration.

Bigrams Revisited

We have seen how the proposed edge-to-letter processing accounts for experimental results on letter perceptibility in non-word strings. Next I show how the proposed letter-to-word processing explains results at the lexical level.

Review of Word-Level Data

Grainger and colleagues have continued their masked form-priming studies into letter-position encoding in experiments in which they systematically varied the positions of target letters included in the prime (Granier et al., 2005; Schoonbaert & Grainger, 2004). In the description of such experiments, the following notation is used for describing the relationship of the prime to the target. A target of length n is represented by 123... n where 1 denotes the first letter, 2 the second letter, etc. The prime is specified in terms of these numbers, with “d” representing a letter not in the target. For example, the prime “rqgdzn” for the target GARDEN is denoted 3d14d6. This means

that the first letter of the prime is the third letter of the target, the second letter of the prime is not in the target, etc.

Positional effects were investigated by using targets of seven and nine letters, where primes consisted of the first four or five letters of the target (initial primes) or the last four or five letters of the target (final primes) (Granier et al., 2005). Control primes consisted entirely of letters not in the target. Within each target length and prime length, initial and final primes provided equivalent levels of facilitation (in terms of speeded lexical-decision RTs). Thus there was no effect of the position of the prime’s letters in the target. These results were taken as evidence of parallel, non-graded activation of bigram units.

However, in another study using primes that contained non-target letters, a positional effect did emerge (Schoonbaert & Grainger, 2004). For seven-letter targets and primes of the form dd34567, 12dd567 or 12345dd, only 12345dd provided facilitation. Thus 12345dd facilitated while dd34567 did not, but 12345 and 34567 provided equivalent levels of facilitation. For five letter strings, there was no effect of position, as neither dd345, 12dd5, nor 123dd produced facilitation, indicating that there was simply not enough overlap with the target (and/or too much overlap with competitors) to provide priming in any of the conditions.

In contrast to the above finding that 12345dd primed while 12dd567 did not, other studies have shown that primes matching on the external letters provided stronger facilitation than those matching on the initial letters (Forster, 1977; Humphreys et al., 1990). Under the SERIOL model, such a pattern could result from an advantage for matching both edge bigrams. The balance between the effects of matching both external letters versus matching initial letters could be influenced by many factors: the degree of overlap with the target, the degree of overlap with competitors, dynamics that vary with prime presentation duration, other experimental parameters, etc. In the following, I will focus on the finding of positional effect for primes with distractor letters versus no positional effect for primes without distractor letters, as these experiments were performed by the same lab under the same methodology, and so are directly comparable.

Another way to investigate positional effects at the word level is to analyze error patterns under degraded perceptual conditions. For very briefly presented word targets (< 50 ms) of six or fewer letters, letters toward the beginning of the target are more likely to be retained in erroneous responses than letters toward the end of the target (Humphreys et al., 1990; Montant et al., 1998). Thus, there is no final-letter advantage in this situation.

A similar pattern is also present in errors made under normal presentation to aphasic patients (Whitney & Berndt, 1999). (Figure 8 shows this data.) A letter in the n th position in the target was scored as being retained if it occurred in n th position in the response. A similar decreasing pattern is also obtained if the n th target letter can occur in any response position, but this pattern is not present if the target and response are aligned at the final letter, and scored backward. Thus there is a strong positional effect that is not an artifact of scoring method. As this pattern is similar to normals under degraded presentation, it is likely that this pattern arises from an inherent aspect of visual word recognition, rather than from brain damage per se.

Thus initial letters have an advantage when noise is present - under aphasia, for very brief presentation to normals, and for primes containing non-target letters. In contrast, for primes consisting only of target letters, no positional effect is evident.

Previous Simulation

In previous work, the aphasic error pattern was replicated in a simulation of the bigram and word levels (Whitney & Berndt, 1999). Bigram activations were taken to depend on both the position and separation of the constituent letters. Aphasia was simulated by adding noise to word activations. Due to the positional dependence of bigram activations, erroneous “responses” tended to preserve the initial letters of the target.

However, this assumption incorrectly predicts that there should be a positional priming effect for initial versus final primes. That is, the reduced activation levels of bigrams toward the end of the word, which allowed replication of the aphasic error pattern, should also yield less priming. However,

as discussed above, this is not the case (Granier et al., 2005).

Another difficulty with the original simulation is that it depended on an additional assumption in order to replicate aphasic error pattern. Normally, bigrams including the final letter were taken to be more activated than bigrams that did not, due to the increased activation of the final letter. However, in order to simulate the finding that the final letter was the least well preserved, it had to be assumed that input to the letter level was reduced, so that there was no final-letter advantage, so that bigrams including the final letter were no more highly activated than those lacking the final letter. Thus, while the bigram-to-word connection weights still reflected the normal advantage for inclusion of the final letter, bigram activations did not. While it is plausible that aphasia would also affect input levels to letter nodes, it would be more desirable to simulate the error pattern without this additional assumption.

Simulation

Therefore, I sought to implement an improved bigram-to-word simulation that demonstrates both a weak positional priming effect, and the strong positional error pattern in the aphasic data, without requiring additional assumptions about activation patterns. In the original simulation, the temporal aspect of bigram and word activations was not considered, nor was lateral inhibition within the word layer. Rather, a bigram vector activated the word layer in a single time step in a purely bottom-up manner. However, a more realistic simulation which includes these factors may allow the above goals to be met. It may be the case that the aphasic error pattern arises from temporal dynamics, rather than from a positional activation pattern. That is, words that are partially activated early (those matching the initial bigrams) could have an advantage over those that are partially activated later, due to ongoing lateral inhibition within the word layer. Therefore, erroneous responses may tend to retain the initial bigrams, even though bigram activations do not vary with position. Based on these ideas, I implemented the following simulation, which met four goals: (1) correct recognition of all

words in a large database; (2) replication of aphasic error pattern under noise; (3) presence of positional effect for primes containing non-target letters; (4) lack of positional effect for primes containing only target letters.

I first give a brief overview of the simulation. The same word database was used as in the original simulation (3650 single-syllable words), and 20 high-frequency 7-letter words were added in order to support the priming simulations. The input layer was comprised of all possible bigram nodes, and the output layer consisted of word nodes representing each word in the database. The input layer connected directly to the output layer. Bigram-to-word weights were set according to the principles specified in the review of the SERIOL model. Bigram activations were clamped sequentially, corresponding to the sequential activation of bigrams. Lateral inhibition within the word layer occurred after each set of bigram activations. Lateral inhibition was included to show that the temporal development of word-level activations could account for the aphasic error pattern and for the positional priming effect under distractor letters. It was not used to simulate settling (reaction) time. Thus the word node having the highest activation following presentation of the final bigram was simply selected as the response.

Next, an example of connection weights and input levels to the word layer is given. Assume that the activation of a bigram node by contiguous letters is 1.0, by a separation of one letter is C_{S1} (< 1.0), and by a separation of two letters is C_{S2} ($< C_{S1}$). For simplicity in this example, we will not consider normalization of connection weights by target length. Table 1 gives the bottom-up input levels to the word node FORM for the stimuli *form* versus *from*.

Next the simulation is specified in more detail. The functions implementing normalization and lateral inhibition were chosen on the basis of convenience and computational efficiency, rather than biological plausibility. I first discuss the bigram activation function and the bigram-to-word weights. As indicated above, some normalization of the weight vector is assumed, to give an advantage for shorter words. In the following, C denotes a parameter. Let B^{xy} denote a bigram node representing the letter x followed by the letter y . Its activation A for a string S is a function of the number of letters

<i>form</i>	Wt	input
*F	1.0	1.0^2
FO	1.0	1.0^2
OR	1.0	1.0^2
FR	C_{S1}	C_{S1}^2
RM	1.0	1.0^2
OM	C_{S1}	C_{S1}^2
FM	C_{S2}	C_{S2}^2
M*	1.0	1.0^2

<i>from</i>	Wt	input
*F	1.0	1.0^2
FR	C_{S1}	$1.0 * C_{S1}$
RO	0.0	$1.0 * 0.0$
FO	1.0	$C_{S1} * 1.0$
OM	C_{S1}	$1.0 * C_{S1}$
RM	1.0	$C_{S1} * 1.0$
FM	C_{S2}	C_{S2}^2
M*	1.0	1.0^2

Table 1: Input levels into the word node FORM, for two different stimuli, *form* (left table) and *from* (right table). In each table, the first column shows the bigrams activated by that stimulus. Bigrams are ordered by their firing sequence, where bigrams within a group fire at the same time. The second column shows the corresponding connection weights into the word node FORM, where each connection weight equals the bigram activation generated by the input *form*. The third column shows the amount of bottom-up input to the word node FORM generated by each activated bigram (*BigramActivation * ConnectionWeight*)

separating x and y . As specified above, $A(B^{xy}, S) = 1.0$ for separation = 0, C_{S1} for separation = 1, C_{S2} for separation = 2, and 0 otherwise. Let Wd^S represent a word node encoding string S . The weight from a bigram node to a word node is given by:

$$W(B^{xy}, Wd^S) = \frac{C_{nrm}}{Len(S) + C_{nrm}} * A(B^{xy}, S)$$

where $Len(S)$ gives the length of the string. This scaling of the bigram's activation value provides normalization by decreasing the weights for longer words, via division by $Len(S)$. The constant C_{nrm} modulates this normalization; the higher its value, the less the effect. (If a bigram received two different activation levels for a word, the larger of $A(B^{xy}, S)$ was taken.)

A string S was presented over $Len(S) + 1$ time steps. At each time step t , the bigrams were clamped to the values that would arise from the activation of the letter in position t . Word-level activations were then updated in two stages. (a) For each word node, the incoming activation was simply added to

the current activation. The incoming activation is given by the dot product of the current bigram vector and the word node’s weight vector. The activations of all word nodes were first updated in this manner. (b) The effects of lateral inhibition were then simulated by re-updating each word node’s activation as follows:

$$A(Wd^S, t) = A(Wd^S, t) - C_{inh} * (AMax(t) - A(Wd^S, t))^2$$

where $AMax(t)$ denotes the activation of the maximally active word node, and the constant $C_{inh}(< 1.0)$ scales the strength of the inhibition.

The parameters were hand-tuned to meet the above four goals. These goals were often at cross purposes. Goal (1) requires normalization of the weight vector. Yet if shorter words have too much of an advantage, they excessively inhibit longer words, under the inhibition required for goals (2) and (3). Goals (2) and (3) require a positional effect, while goal (4) requires no positional effect.

A range of parameter values near the following values yielded reasonable results; the results for these particular values are presented.

$$C_{S1} = 0.8 \quad C_{S2} = 0.4 \quad C_{nrm} = 20 \quad C_{inh} = 0.1$$

All words in the database were recognized correctly, under the requirement that difference between the activation of the target word and the next highest word be at least 0.2. The most challenging task was to distinguish between *the*, *tee*, *thee*, *teeth*, and *teethe*.

Priming

Five priming conditions were simulated for seven-letter targets: 12345, 34567, 12345dd, 12dd567, and 12345dd. The seven-letter words added to the database were used as targets. A prime string was generated for each target for each prime condition. Priming was simulated by presenting the prime string as input, and then reading off the activation of the corresponding target word node. For example, in a trial simulating the 12345dd condition for the target

Prime	Act.
12345	5.3
34567	4.9
12345dd	5.3
12dd567	3.1
dd12345	3.0

Table 2: Mean activation of target nodes in priming simulation.

without, the string *withoak* was encoded over the bigram nodes. Following activation of the final edge bigram (K^*), the activation of the word node WITHOUT was recorded.

To form the prime strings for the dd34567 and 12dd567 conditions, letter replacements were primarily chosen from letter pairs in the alphabet (frequently yielding illegal letter combinations), and the distractor letters did not form actual words in conjunction with nearby target letters. For 12345dd, the distractor letters were chosen such that 5dd formed a word in the database in most cases (such as *oak* in the *withoak* example). This choice of distractor letters follows the observation (Davis, pers. comm.) that distractor / target letter combinations in the experimental stimuli (Schoonbaert & Grainger, 2004) were more similar to actual words in the 12345dd condition than the dd34567 condition. The mean target activation levels for the five priming conditions are given in Table 2.

It is evident that 12345, 34567, and 12345dd yielded the highest activation levels, consistent with the experimental findings that those primes produced facilitation, while 12dd567 and dd34567 did not. In the simulation, the latter primes yielded lower activation levels because of the early inhibition from other word nodes activated by the distractor letters. Although the distractor letters did not actually form words in conjunction with the target letters, the distractor letters nevertheless caused non-target words to become partially activated, due to the flexibility of the bigram encoding. The inhibition from these competitors then put the target word at a disadvantage from which it could not recover.

Despite the fact that 12345dd primes contained actual words, target word-

node activations were high because the target word got an early advantage, and inhibited the competitors. Thus the simulation results for replacement primes depend directly on the temporal sequence of bigram activations, going in the opposite direction of what would be expected on the basis of similarity to competitor words, in agreement with the experimental results.

For the primes containing no distractor letters, inhibition from competitors was reduced, and the difference between initial and final primes was small. This small difference is numerically consistent with the results of Granier et al. (2005), where a non-significant advantage (< 10 ms) for initial primes over final primes was consistently observed across five experiments.

Aphasic Error Pattern

In evaluating the experimental data on errors made by aphasics, a letter was counted as being retained if occurred in the erroneous response in the same position as in the target word (Whitney & Berndt, 1999). More complex methods of scoring yielded the same positional patterns as this position-specific method, so the position-specific metric was used for simplicity. In order to compare the simulated data to the experimental data, I use the same metric.

A lesion was simulated by adding normally distributed noise to each word node at each time step (prior to the inhibition). Noise with mean 0.3 and standard deviation 0.45 yielded good results, shown in figure 8. As is evident, the probability of retaining a letter decreased with its position (in particular, across positions 1 through 4). In contrast, backward scoring of the results gave a much flatter pattern (as was also the case for the experimental data (Whitney & Berndt, 1999)), indicating that the forward positional effect was not merely an artifact of the scoring method. In the simulation, this positional effect arose from the sequential activation of the bigram nodes, coupled with the ongoing lateral inhibition. Potential erroneous responses that were not initially highly activated became inhibited and remained at a disadvantage. Therefore, retention level was highest for early string positions and decreased across the string, giving a strong positional effect.

For comparison, the same simulation was also run without lateral inhibition. The results are given in Figure 9. In this case, forward and backward scoring yielded the same pattern; a weaker positional effect is evident in both cases. These positional effects are therefore an artifact of the position-specific scoring method. (As distance from an edge letter increases, the probability of position-altering insertions and deletions increases, reducing the probability of a retained letter remaining in the same position as in the target.)

The results of the lesioned simulation also showed other similarities to the experimental data. Aphasic subjects tended to preserve word length in their erroneous responses. Average response length to targets of lengths 3 to 6 were 4.0, 4.2, 4.9, and 5.9, respectively (Whitney & Berndt, 1999). The simulated data also showed sensitivity to target length, giving 4.3, 4.6, 5.0, and 5.6. Retention level at a given position tended to increase with target length for both the aphasics and the simulation. For example, for position 3, experimental retention rates were 40%, 55%, 65%, and 55% for targets of lengths 3 to 6, respectively, (Whitney & Berndt, 1999), while the simulation gave 23%, 40%, 54%, and 62%.

Thus the simulation accomplished the stated goals. There was a very weak positional effect for deletion primes, and strong positional effects for primes containing distractor letters, and for erroneous responses in the aphasic simulation. Under the addition of noise (at the word level or in the form of distractor letters), the effects of lateral inhibition were accentuated, and a positional influence was revealed which resulted from the serial nature of bigram activations, without any additional assumptions about letter activation patterns. In contrast, for primes that only contained target letters, the absolute level of bottom-up input from the bigram nodes dominated, and there was little effect of position of match. The underlying principles of the simulation explain this conflicting data on positional effects at the word level.

Discussion

We have seen that there are no data that actually contradict a serial encoding, and have considered how the serial encoding proposed in the SERIOL model

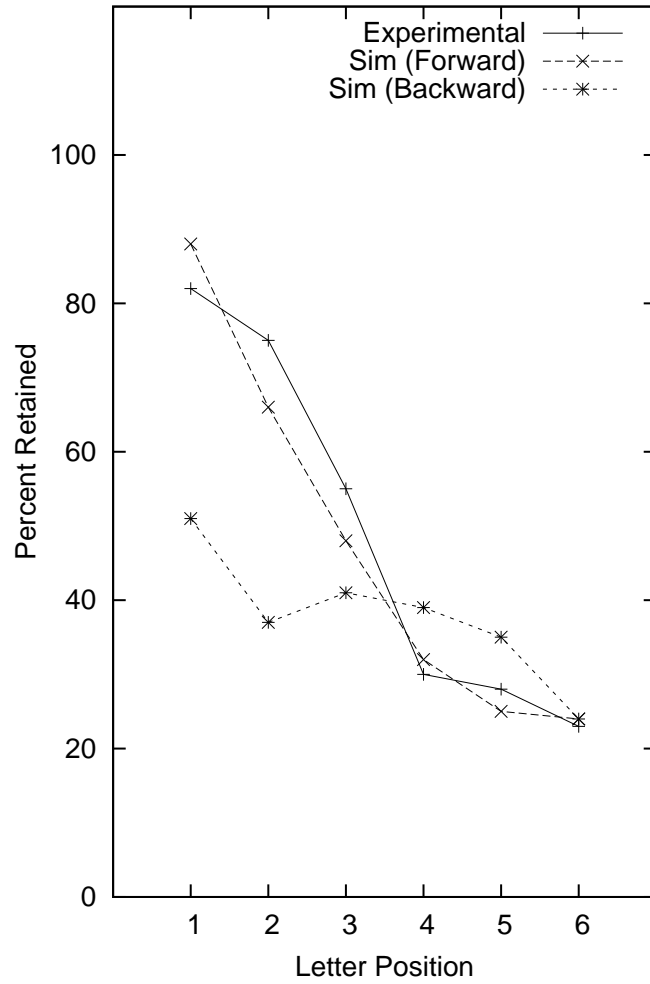


Figure 8: Experimental (Whitney & Berndt, 1999) and simulated results for the aphasic error pattern, with backward scoring of simulated results for comparison. *Percent Retained* refers to the percentage of erroneous trials in which the letter in the i th position in the target occurred in the i th position the response ($n = 201$ for experiment; $n = 367$ for simulation). In backward scoring, the target and response were aligned at the final letter, and scored from right to left. In this case, position 1 corresponds to the final letter, 2 corresponds to the next-to-last letter, etc. Data are collapsed over target lengths of three to six. (In the both the experimental data and the simulation, there same patterns were present within each target length.)

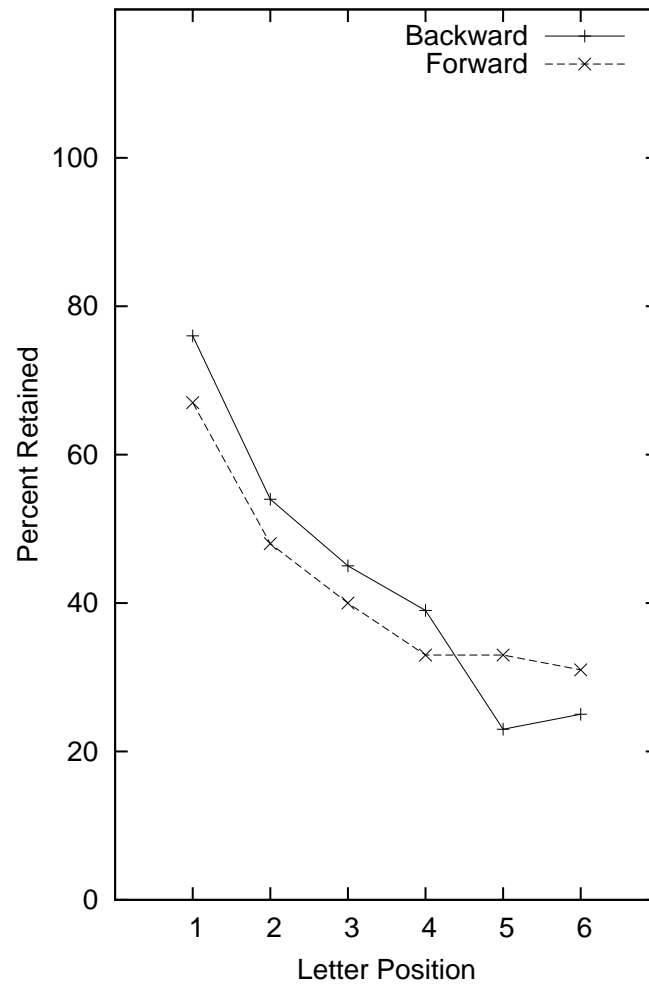


Figure 9: Results of simulation without lateral inhibition, for forward versus backward scoring.

provides important representational capacities. This pre-existing model allows accounts of experimental data that are otherwise difficult to explain. In particular, the explanation of the final-letter primacy in the RVF (and its absence for very brief exposures) crucially depends on the proposed serial-encoding mechanism. I note that the SERIOL model was not designed to explain this pattern. Rather, the explanation falls out from the known shape of the acuity gradient, the specification of how the locational gradient is formed, and the interaction of the locational gradient with the oscillatory cycle. As discussed above, an asymmetry in the strength of lateral masking cannot account for these results.

Another alternative explanation is that the final letter is perceived best in the RVF because final letters usually fall in the RVF (and initial letters fall in the LVF). However, this account is insufficient. First of all, it does not provide an explicit mechanism. Second, it can't explain the dependence of the final-letter primacy on exposure duration. Third, the underlying assumption is incorrect. Many short, high-frequency words are not directly fixated during naturalistic reading. Rather they are completely processed during parafoveal preview (Rayner, 1975). Thus, initial letters often also fall in the RVF. In contrast, the SERIOL model accounts for these counterintuitive perceptibility findings in a precise, thorough manner.

Although it may be suggested that letter perceptibility patterns are not necessarily germane to the issue of letter-position encoding and lexical access, the data indicate otherwise. For central fixation, external symbol perceptibility differs for letter and non-letter strings (Hammond & Green, 1982; Mason, 1982). For unilateral presentation of letter strings, accuracy patterns are contrary to the expected effect of acuity (Bouma, 1973; Estes, Allemeyer & Reder, 1976; Legge, Mansfield & Chung, 2001). Furthermore, perceptibility patterns differ with reading direction for identification of a single letter within a string of homogeneous distractor letters (Nazir, Ben-Boutayab, Decoppet, Deutsch & Frost, 2004). In this case, the target letter could “pop out”, so string processing is not required. Yet, reading direction affected the results, indicating a low-level influence of string-processing mechanisms. I maintain that these experimental results provide strong evidence for string-

specific processing, and also provide important clues as to the nature of that processing. Therefore, I suggest that any model of letter-position encoding should account for these patterns.

We have also seen that serial lexical access, in conjunction with ongoing lateral inhibition within the word level, explains positional effects in the presence of noise (i.e., non-target letters, brain damage, or very brief presentation). If a non-target is activated late in the oscillatory cycle, it is inhibited strongly by the more active word nodes, thereby suffering a competitive disadvantage. Thus there is an advantage for matching the initial letters of the target, yielding an positional effect. When a prime consists only of target letters, competition is less relevant and a positional effect does not emerge, because bigram activation levels are not influenced by position. In contrast, the proposal of parallel lexical access cannot explain these differing patterns.

Previously, the pendulum swung from the assumption of a serial readout of a letter string (Gough, 1972; Harcum & Nice, 1975; Mewhort, Merikle & Bryden, 1969; Sperling, 1963) to the assumption of parallel processing (Coltheart et al., 2001; McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989). Given the lack of a valid argument against a serial encoding, the direct evidence for serial processing (Nice & Harcum, 1976), the ability of a serial encoding to provide important representational functions, and the explanatory capacity of SERIOL's serial mechanism, I suggest that it is time for the pendulum to swing back to seriality (and remain there).

References

- Adrian, E. (1928). *The Basis of Sensation*. New York, NY: Norton.
- Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin and Review*, 4, 439–461.
- Balota, D., Cortese, M., Sergent-Marshall, S. & Spieler, D.H. Yap, M. (2004). The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords. <http://elexicon.wustl.edu>. Washington University.
- Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, 13, 767–782.
- Burns, B. (1968). *The Uncertain Nervous System*. London, UK: Arnold.
- Chastain, G. (1989). Effects of relative position and shape relation on gap detection with geometric form pairs. *Acta Psychologica*, 72, 13–23.
- Coltheart, M., Davelaar, E., Jonasson, J. & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535–555). Hillsdale, New Jersey: Lawrence Earlbaum Associates.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–256.
- Davis, C. (1999). *The Self-Organising Lexical Acquisition and Recognition (SOLAR) model of visual word recognition*. PhD thesis, University of South Wales.
- Ellis, A. W., Young, A. W. & Anderson, C. (1988). Modes of word recognition in the left and right cerebral hemispheres. *Brain and Language*, 35, 254–273.

- Estes, W., Altemeyer, D. & Reder, S. (1976). Serial position functions for letter identification at brief and extended exposure durations. *Perception and Psychophysics*, 19, 1–15.
- Forster, K. (1977). Accessing the mental lexicon. In R. Wales & E. Walker (Eds.), *New Approaches to Language Mechanisms* (pp. 535–555). Amsterdam: New Holland.
- Frederiksen, J. & Kroll, J. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 361–379.
- Friedmann, N. & Gvion, A. (2001). Letter position dyslexia. *Cognitive Neuropsychology*, 18, 673–696.
- Goodman, J. C. & Huttenlocher, J. (1988). Do we know how people identify spoken words? *Journal of Memory and Language*, 27, 684–698.
- Gough, P. (1972). One second of reading. In J. Kavanagh & I. Mattingly (Eds.), *Language By Ear and by Eye: The Relationships between Speech and Reading* (pp. 331–358). Cambridge: MIT Press.
- Grainger, J. & Whitney, C. (2004). Does the human mind read words as a whole? *Trends in Cognitive Science*, 8, 58–59.
- Granier, J., Grainger, J., Farioli, Q. & van Heuven, W. J. (2005). Letter position coding in printed word perception: Evidence from relative-position priming. In Preparation.
- Hammond, E. & Green, D. (1982). Detecting targets in letter and non-letter arrays. *Canadian Journal of Psychology*, 36, 67–82.
- Harcum, E. & Nice, D. (1975). Serial processing shown by mutual masking of icons. *Perceptual and Motor Skills*, 40, 399–408.
- Harm, M. W. & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review*, 106, 491–528.

- Helenius, P., Tarkiainen, A., Cornelissen, P., Hansen, P. C. & Salmelin, R. (1999). Dissociation of normal feature analysis and deficient processing of letter-strings in dyslexic adults. *Cerebral Cortex*, 9, 476–483.
- Hopfield, J. J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature*, 376, 33–36.
- Humphreys, G. W., Evett, L. J. & Quinlan, P. T. (1990). Orthographic processing in visual word identification. *Cognitive Psychology*, 22, 517–560.
- Kwantes, P. J. & Mewhort, D. J. K. (1999). Evidence for sequential processing in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 376–381.
- Lamberts, K. (2005). Interpretation of orthographic uniqueness point effects in visual word recognition. *Journal of Experimental Psychology: Human Perception And Performance*, 31, 14–19.
- Lee, H., Simpson, G., Logothetis, N. & G., R. (2005). Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron*, 45, 147–156.
- Lefton, L., Fisher, D. & Kuhn, D. (1978). Left-to-right processing of alphabetic material is independent of retinal location. *Bulletin of the Psychonomic Society*, 12, 171–174.
- Legge, G. E., Mansfield, J. S. & Chung, S. T. (2001). Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research*, 41, 725–743.
- Lindell, A., Nicholls, M. & Castles, A. (2003). The effect of orthographic uniqueness and deviation points on lexical decisions: Evidence from unilateral and bilateral-redundant presentations. *Quarterly Journal of Experimental Psychology*, 56, 287–307.
- Lisman, J. E. & Idiart, M. A. (1995). Storage of 7 +/-2 short-term memories in oscillatory subcycles. *Science*, 267, 1512–1515.

- Mainen, Z. & Sejnowski, T. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268, 1503–1506.
- Marslen-Wilson, W. D. & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29–63.
- Mason, M. (1982). Recognition time for letters and non-letters: Effects of serial position, array size, and processing order. *Journal of Experimental Psychology*, 8, 724–738.
- McClelland, J. L. & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375–407.
- Mewhort, D., Merikle, P. & Bryden, M. (1969). On the transfer from iconic to short-term memory. *Journal of Experimental Psychology*, 81, 89–94.
- Montant, M., Nazir, T. A. & Poncet, M. (1998). Pure alexia and the viewing position effect in printed words. *Journal of Cognitive Neuropsychology*, 15, 93–140.
- Mozer, M. (1991). *The Perception of Multiple Objects: A Connectionist Approach*. Cambridge: MIT Press.
- Nazir, T. A., Ben-Boutayab, N., Decoppet, N., Deutsch, A. & Frost, R. (2004). Reading habits, perceptual learning, and recognition of printed words. *Brain and Language*, 88, 294–311.
- New, B., Ferrand, L., Pallier, C. & Brysbaert, M. (2005). Re-examining word length effects in visual word recognition: New evidence from the English Lexicon Project. Submitted.
- Nice, D. & Harcum, E. (1976). Evidence from mutual masking for serial processing of tachistoscopic letter patterns. *Perceptual and Motor Skills*, 42, 991–1003.

- Perea, M. & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory and Cognition*, 31, 829–841.
- Peressotti, F. & Grainger, J. (1995). Letter-position coding in random constant arrays. *Perception and Psychophysics*, 57, 875–890.
- Peressotti, F. & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception and Psychophysics*, 61, 691–706.
- Perry, C. & Ziegler, J. C. (2002). Cross-language computational investigation of the length effect in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 990–1001.
- Radeau, M., Morais, J., Mousty, P., Saerens, M. & Bertelson, P. (1992). A listener’s investigation of printed word processing. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 861–871.
- Radeau, M., Mousty, P. & Bertelson, P. (1989). The effect of the uniqueness point in spoken-word recognition. *Psychological Research*, 51, 123–128.
- Rayner, K. (1975). Parafoveal identification during a fixation in reading. *Acta Psychologica*, 39, 271–282.
- Rayner, K. & Pollatsek, A. (1989). *The Psychology of Reading*. Englewood Cliffs, NJ: Prentice-Hall.
- Reicher, G. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81, 275–280.
- Richardson, J. (1976). The effects of stimulus attributes on latency of word recognition. *British Journal of Psychology*, 67, 315–325.
- Rieke, F., Warland, D., de Ruyter von Stevenink, R. & Bialek, W. (1997). *Spikes: Exploring the Neural Code*. Cambridge, MA: MIT Press.

- Roberts, M. A., Rastle, K., Coltheart, M. & Besner, D. (2003). When parallel processing in visual word recognition is not enough: new evidence from naming. *Psychonomic Bulletin Review*, 10, 405–414.
- Rumelhart, D. & McClelland, J. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volumes 1 and 2*. Cambridge, MA: MIT Press.
- Rumelhart, D. E. & McClelland, J. L. (1982). 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.
- Schoonbaert, S. & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, 19, 333–367.
- Seidenberg, M. S. & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.
- Seidenberg, M. S., Plaut, D. C., Petersen, A. S., McClelland, J. L. & McRae, K. (1994). Nonword pronunciation and models of word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1177–1196.
- Sperling, G. (1963). A model for vision memory tasks. *Human Factors*, 5, 19–31.
- de Ruyter van Stevenick, R., Lewen, G., Strong, S., Koberle, R. & Bialek, W. (1997). Reproducibility and Variability in Spike Trains. *Science*, 275, 1805–1808.
- Tesche, C. & Karhu, J. (2000). Theta oscillations index human hippocampal activation during a working memory task. *Proceedings of the National Academy of Sciences*, 97, 919–924.

- VanRullen, R. & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences*, 7, 207–213.
- Ward, L. (2003). Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences*, 7, 553–559.
- Westheimer, G. (1987). Visual acuity. In R. Moses & W. Hart (Eds.), *Adler's Physiology of the Eye: Clinical Application* (pp. 531–547). St. Louis, MO: The C.V. Mosby Company.
- Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin and Review*, 8, 221–243.
- Whitney, C. & Berndt, R. S. (1999). A new model of letter string encoding: Simulating right neglect dyslexia. *Progress in Brain Research*, 121, 143–163.
- Whitney, C. & Cornelissen, P. (2005). Letter-position encoding and dyslexia. *Journal of Research in Reading*, 28, 274–301.
- Whitney, C. & Lavidor, M. (2004). Why word length only matters in the left visual field. *Neuropsychologia*, 42, 1680–1688.
- Whitney, C. & Lavidor, M. (2005). Facilitative orthographic neighborhood effects: The SERIOL model account. *Cognitive Psychology*. In Press.
- Wolford, G. & Hollingsworth, S. (1974). Retinal location and string position as important variables in visual information processing. *Perception and Psychophysics*, 16, 437–442.
- Young, A. & Ellis, A. (1985). Different methods of lexical access for words presented to the left and right visual fields. *Brain and Language*, 24, 326–358.