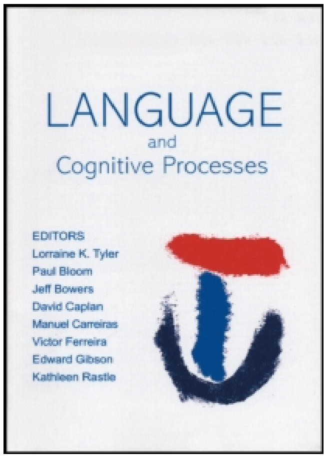


Language and Cognitive Processes

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683153

SERIOL Reading

Carol Whitney [a]; Piers Cornelissen [b]
[a] Department of Linguistics, University of Maryland, College Park, MD, USA
[b] Department of Psychology, University of York, York, UK

First Published on: 07 November 2007
To cite this Article: Whitney, Carol and Cornelissen, Piers (2007) 'SERIOL Reading', Language and Cognitive Processes, 23:1, 143 - 164
To link to this article: DOI: 10.1080/01690960701579771
URL: http://dx.doi.org/10.1080/01690960701579771

LANGUAGE AND COGNITIVE PROCESSES
2008, 23 (1), 143–164

Psychology Press
Taylor & Francis Group

SERIOL Reading

Carol Whitney
Department of Linguistics, University of Maryland, College Park, MD, USA

Piers Cornelissen
Department of Psychology, University of York, York, UK

In recent years, there has been a growing interest in the issue of orthographic processing in visual word recognition. We compare the SERIOL model to parallel models in their ability to explain experimental data at the lexical and sub-lexical levels. We find that the parallel models have difficulty accounting for position-independent letter priming, letter perceptibility patterns, and temporal effects.

INTRODUCTION

Following the publication of the SERIOL model of letter-position encoding (Whitney, 2001a), interest in understanding orthographic processing has grown and several alternative models have been presented. The overall goal of this paper is to compare the SERIOL model to other models in their ability to explain experimental results. A commentary on one competing model, SOLAR (Davis, 1999; Davis & Bowers, 2006), is presented elsewhere (Whitney, in press). Here, we concentrate on a group of models that include multi-letter units and purely parallel processing. (Dehaene, Cohen, Sigman, & Vinckier, 2005; Grainger, Granier, Farioli, Van Assche, & van Heuven, 2006; Grainger & van Heuven, 2003). We also address specific criticisms of the SERIOL model (Dehaene et al., 2005; Goswami & Ziegler, 2006).

The organisation of this article is as follows. We start by defining the problem of letter-position encoding, and discussing its importance. Next, we briefly review the four models under consideration. We then compare the models in various contexts. First, we review letter-priming experiments and

---

Correspondence should be addressed to: Carol Whitney, Department of Linguistics, University of Maryland, College Park, MD 20740, USA. E-mail: cwhitney@cs.umd.edu

http://www.psypress.com/lcp DOI: 10.1080/01690960701579771

discuss whether the models can account for these data. Second, we discuss the models' ability to explain letter perceptibility patterns. Third, we discuss serial processing. Fourth and fifth, we consider the robustness of the proposed representations for direct lexical access and for sub-lexical phonological processing.

LETTER-POSITION ENCODING

When a visual image first arrives at the cortex, it is in the form of a retinotopic encoding. If the stimulus is a string, this spatial representation must be transformed to an encoding of the constituent letters' identities and positions, to provide a suitable representation for lexical access. There is a growing consensus that reading involves multiple processing routes. On the lexical route, a string directly accesses lexical representations. On the sub-lexical route, a string is mapped onto a phonological representation, which then contacts lexical representations. Thus one aspect of the problem of letter-position encoding is to define how a retinotopic representation is converted into abstract encoding of letter order, which provides input to the lexical and sub-lexical routes. A further aspect is to define how this encoding then activates lexical representations along the lexical route in particular. Note that the problem of letter recognition itself is not addressed. Rather, the question is more abstract – given the ability to identify letters, how are the relationships between letters derived from a spatial encoding, and represented to allow lexical and phonological processing?

This question is highly relevant to the problem of developmental dyslexia. Although it is widely assumed that dyslexia stems from a core phonological deficit, this is not necessarily the case (Castles & Coltheart, 2004). There is evidence that aspects of visual processing are abnormal in many dyslexics (e.g., Buchholz & Aimola Davies, 2006; Cornelissen, Richardson, Mason, & Stein, 1995; Facoetti & Molteni, 2001; Livingstone, Rosen, Drislane, & Galaburda, 1991; Sperling, Zhong-lin, Manis, & Seidenberg, 2003). If the visual system cannot provide an accurate encoding of letter order, this will interfere with the ability to learn the relationships between letters and sounds, leading to an impaired ability to read pseudowords and to perform phonological-awareness tasks (Whitney & Cornelissen, 2005). Thus a visual deficit could cause difficulties in phonological processing, even in the absence of a core phonological deficit. In order to understand what could go awry in dyslexics' visual processing, it is first necessary to understand normal visual processing in reading.

The question of letter-position encoding is also relevant to the field of visual object recognition in general. An abstract orthographic encoding entails a representation that is not tied to retinal location, but which encodes spatial relationships between the letters. The problem of representing sub-part relationships in a location-invariant manner is a basic issue in visual object recognition. Understanding how the brain solves this problem in the restricted domain of visual word recognition may shed light on general mechanisms of visual processing.

Thus, understanding how letter position is encoded is of great practical and theoretical significance. Therefore, it is important to identify an accurate model of orthographic processing. We next review four contenders, and then evaluate them on their representational and explanatory capacity.

MODEL REVIEWS

Figure 1 summarises the models under consideration. The models will each be presented in a bottom-up manner. For brevity, we will simply specify the proposed representations and not address the motivations for choosing them. For the SERIOL model, these issues are addressed in detail elsewhere (Whitney & Berndt, 1999; Whitney, 2001a). For clarity, capitalised boldface will be used to denote a model unit encoding an item, while italics will be

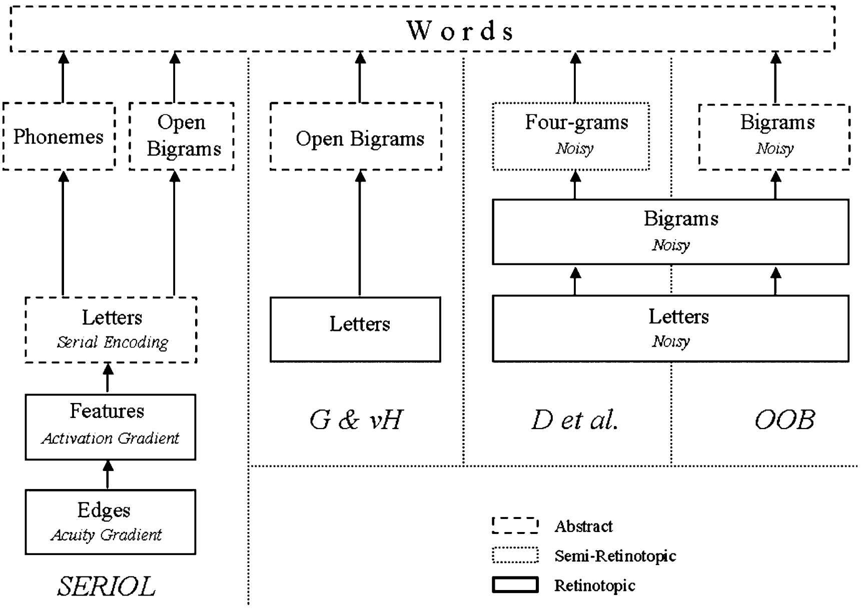


Figure 1. Overview of the four models under consideration, from left to right: the SERIOL model (Whitney, 2001a, 2004a); the model of Grainger & van Heuven (2003); the model of Deheane et al. (2005); the Overlap Open-Bigram model (Grainger et al., 2006)

used for a stimulus. For example, **CART** denotes a word unit that recognises the string *cart*.

SERIOL

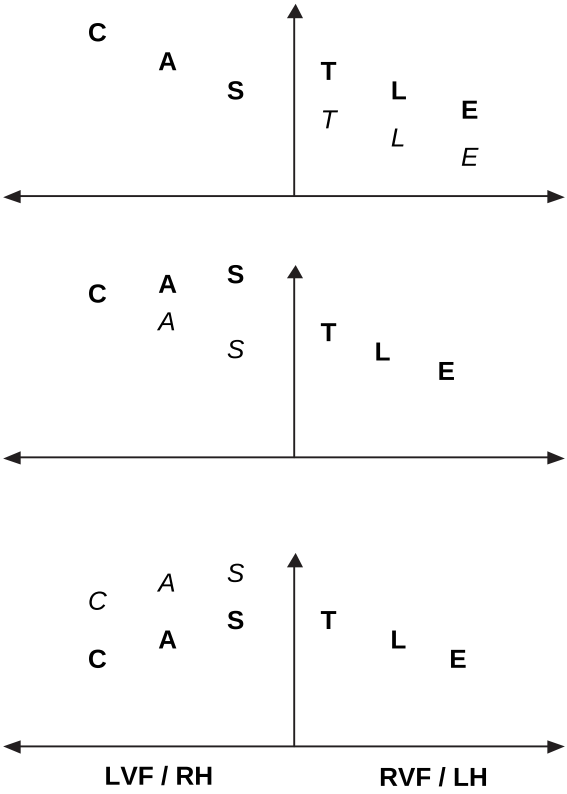
The SERIOL model is a theoretical framework that describes how information is represented at each of its processing layers and specifies how each representation is transformed into the encoding at the next layer. It is summarised in Figure 1. The model focuses on bottom-up processing, but this is not meant to rule out top-down interactions.

The SERIOL model is comprised of five layers: edges, features, letters, open-bigrams, and words. The first two layers are retinotopic, while the latter three are abstract. For the retinotopic layers, the term *activation level* will be used to denote the total amount of neural activity devoted to representing a letter within a given layer. A letter's activation level increases with the number of neurons representing that letter and their firing rate. For the abstract layers, the term *activation* will denote the activity level of a representational unit in a given layer.

The edge layer models V1/V2. In these areas, the rate of spatial sampling (acuity) is known to sharply decrease with increasing eccentricity. This is modelled by the assumption that activation level decreases as distance from fixation increases. This pattern is termed the *acuity gradient*.

The feature layer models V4. Based on learned, hemisphere-specific processing, the acuity gradient of the edge layer is converted to a monotonically decreasing activation gradient (dubbed the *locational gradient*) in the feature layer. That is, activation level is highest for the first letter, and decreases across the string. Hemisphere-specific processing is necessary because, for a fixated word, the acuity gradient does not match the locational gradient in the first half of the word (i.e., acuity *increases* from the first letter to the fixated letter, whereas the locational gradient *decreases* across the string), whereas the acuity gradient and locational gradient match in the second half of the word (i.e., both decreasing). Strong directional lateral inhibition is required in the hemisphere contralateral to the first half of the word, in order to invert the acuity gradient. This processing is summarised in Figure 2.

At the letter layer, corresponding to posterior fusiform gyrus, letter units fire serially. That is, the letter unit encoding the first letter fires, then the unit encoding the second letter fires, etc. This mechanism is based on the general proposal that item order is encoded in successive gamma cycles (60 Hz) of a theta cycle (5 Hz) (Lisman & Idiart, 1995). That is, each activated letter unit fires in a burst for about 15 ms (one gamma cycle), and bursting repeats every 200 ms (one theta cycle). Activated letter units burst slightly out of phase with each other, such that they fire in a rapid sequence. This firing



RH inhibits LH.

Left-to-right in hibition in RH.

Stronger excitation in RH..

Figure 2. Illustration of the formation of the locational gradient at the feature level for the centrally fixated stimulus *castle*. The vertical axis represents activation level, and the horizontal axis represents retinotopic space, with fixation at the origin. The processing is comprised of three transformations. For clarity, these transformations are shown as occurring sequentially, although they would actually occur interactively. In each panel, the boldface letters represent the activation pattern prior to the illustrated transformation, and the italics represent the result of the transformation. Bottom Panel: the bold letters correspond to the acuity gradient. Stronger RH excitation brings the first letter (as well as the other RH letters) to a higher activation level. Middle Panel: the first letter inhibits the second letter, and the first two letters inhibit the third. Thus RH, left-to-right inhibition inverts the acuity gradient. Top Panel: cross-hemispheric inhibition 'joins' the hemispheric gradients. The final result (bold letters in LVF/RH and italicised letters in RVF/LH) is a monotonically decreasing activation gradient

pattern is induced by the interaction of the locational gradient with subthreshold theta oscillations within the letter units; lateral inhibition between letter units helps to maintain seriality. This mechanism also causes letter units to have varying activations (i.e., different spiking rates and burst durations), where activation generally decreases across the string, but rises for the final letter. See Whitney (2001a) for details. Note that the serial encoding is the key point of abstraction; the retinotopic representation is mapped onto a temporal one, providing a location-invariant representation of letter order.

This abstract, serial encoding provides input to both the lexical and sub-lexical routes. It is assumed that the sub-lexical route parses and translates the sequence of letters into a grapho-phonological encoding (Whitney & Cornelissen, 2005). That is, the resulting representation encodes syllabic structure and records which graphemes generated which phonemes. However, a more detailed description of the neural underpinnings of this processing is currently beyond the scope of the model. The remaining layers of the model address processing that is specific to the lexical route.

At the open-bigram (Grainger & Whitney, 2004) layer, corresponding to left middle fusiform, units recognise pairs of letter units that fire in a particular order. That is, open-bigram unit **XY** is activated when letter-unit **X** fires before **Y**, where the letters *x* and *y* were not necessarily contiguous in the string. The activation of an open-bigram unit decreases with increasing time between the firing of the constituent letter units. The open-bigram units connect to the word layer, via weighted connections. The input to a word unit is the dot-product of the bigram vector and its weight vector.

Whitney (2004a) specified several refinements to the open-bigram representation first presented in Whitney and Berndt (1999). (1) Originally, open-bigram activations also depended on the constituent letter units' activations (which depend on string position). This assumption has been dropped; bigram activations depend only on the distance between the constituent letters. Experimental data that were originally explained by positional variations in bigram activations are now explained directly by seriality (Whitney, 2004a, 2007). (2) The representation is now taken to include edge units, which explicitly encode the first and last letters. For example, the encoding of the stimulus CART would be ***C**, **CA**, **AR**, **CR**, **RT**, **AT**, **CT**, and **T***, where * represents an edge. (3) Normalisation of weights is assumed, such that weights to shorter words are larger than weights to longer words. For example, the connection weights from **CA**, **AN**, and **CN** to the word-unit **CAN** are larger than the weights to **CANON**. Hence, the stimulus *can* would activate **CAN** more than **CANON**. Changes (2) and (3) replace the original assumption that the bigram encoding the external letters possessed special properties.

Grainger and van Heuven (2003)

In this model, an alphabetic array is converted in parallel into an open-bigram encoding, where open-bigram activations are either 1 or 0. Open bigrams are activated by letter pairs having up to two intervening letters. The open-bigram units then connect to word units.

The alphabetic array is a retinotopic encoding, in which a given letter has a separate representation for each retinal location. Thus the conversion from the alphabetic array to the open-bigram representation creates a location invariant representation. However, the model does not specify the underlying mechanisms of this conversion.

Dehaene et al. (2005)

This model starts with noisy retinotopic letter array. That is, the occurrence of a letter in given retinal location activates the representation of that letter in the corresponding location in the array, and partially activates representations of that letter in nearby locations in the array. At the next layer of processing, retinotopic detectors respond to bigrams (ordered *contiguous* letter pairs), rather than to open-bigrams. However, due to the noise in the letter level, some of the activated bigram detectors represent open-bigrams and transpositions. The next level consists of detectors for ordered contiguous pairs of bigrams (i.e., four-grams), which connect to word units.

The gist of the model is that location-invariance is achieved by gradually increasing receptive-field size. However, the four-grams would also have to be retinotopic (in order to recognise open-bigrams in a particular order), so the model does not actually include a location-invariant encoding of letter order. For example, under the given scheme, learning a novel four-letter word presented at fixation would not allow recognition of that word some distance from fixation.

Overlap Open-Bigram model

Grainger et al. (2006) presented a variation of the above models, dubbed the Overlap Open-Bigram model (OOB), in which the first two layers of the model are the similar to those of Dehaene et al. (2005) (i.e., noisy retinotopic letters and bigrams). The retinotopic bigrams then activate abstract bigrams, which activate the word level. The graded activity levels at the letter layer (full versus partial activation) propagate to the bigram levels. As a result, bigram detectors representing contiguous bigrams are more highly activated than those encoding open-bigrams or transpositions.

Thus, this model improves on the above two models, in that it includes a location-invariant encoding and a mechanism for bigram activation. Within the bigram level, the OOB model is similar to the SERIOL model in that it employs graded bigram activations, except that the OOB model also includes activation of bigrams corresponding to transpositions.

Summary

The models differ in three key ways as illustrated in Figure 1.

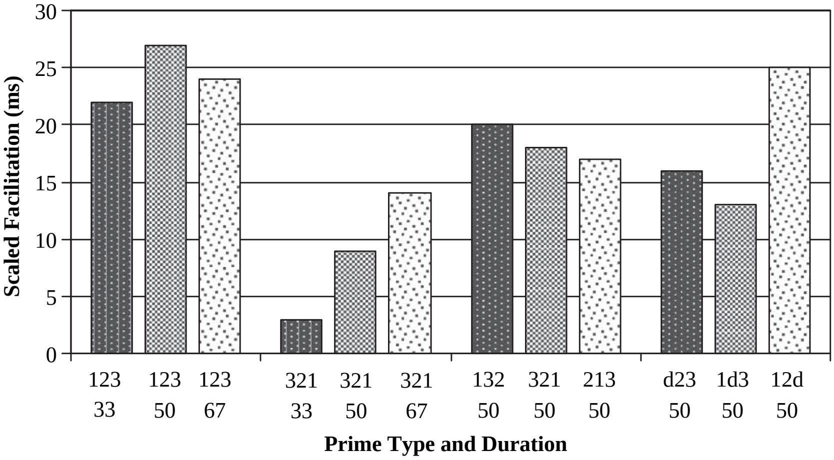
1. Abstraction occurs at different levels of processing. The SERIOL model is the only one that specifies an abstract representation of individual letters. Such a letter unit can represent that letter in any retinal location, wherein timing firing binds positional information to the letter identity. In contrast, there are separate letter units for a given letter at different retinal locations in the other models.
2. The SERIOL model includes a serial encoding, while the other models are purely parallel. In the SERIOL model, space is mapped onto time to create an abstract, invariant representation. This abstraction mechanism allows an invariant representation to be achieved at a lower level of processing (i.e., individual letters) than the other models.
3. The SERIOL model specifies processing below the letter level, while the other models do not. As argued in the upcoming sections, the proposed processing accounts for a variety of orthographic effects that the other models do not address.

In comparing the models, we will first look at experimental results on non-pronounceable letter strings. Such stimuli isolate processing at a pre-lexical, orthographic, non-phonological level. Priming experiments indicate the existence of abstract letter representations, while perceptual studies show patterns specific to alphanumeric status, visual field, and reading direction.

LETTER PRIMING

Peressotti and Grainger (1995) performed a series of alphabetic-decision experiments that are quite informative as to the nature of orthographic encoding at the letter level. The task was to identify whether a three-character stimulus consists entirely of letters. Targets were comprised of three consonants, while foils consisted of two consonants and a non-letter character, such as ‘&’. A briefly presented prime (33, 50, or 67 ms) and a mask preceded the stimulus. The goal was to investigate the effect of different relationships between prime and target. To indicate such relationships, the prime’s letters will be denoted by their position in the target, with ‘d’ indicating a letter not in the target. For example, for the target GDK, the prime GDK is denoted 123, KGD is 312, and LDK is d23.

The results of the experiments are summarised in the upper panel of Figure 3. Several things are evident. For same-position primes (123), facilitation is robust at all exposure durations. The other prime types give



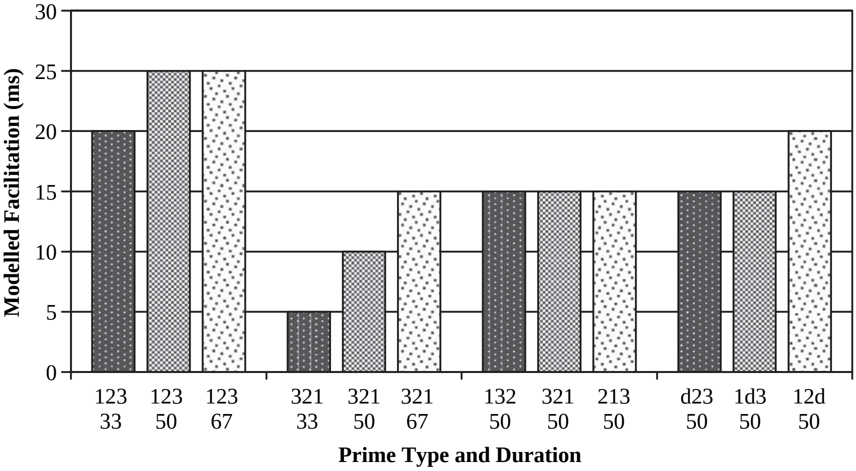


Figure 3. Upper Panel: Summary of experimental results from Peressotti and Grainger (1995). For same-position and cross-position primes (123 and 312), presentation durations were varied. For transposition (132, 321, 213) and replacement (d23, 1d3, 12d) primes, prime type was varied while presentation duration was held constant. Facilitation figures for the first three groups (same-position, cross-position, and transposition primes) directly reflect the experimental findings, which were obtained under a go/no-go procedure. The replacement primes were run under a yes/no procedure, which gave reduced facilitation overall. In order to compare replacement primes to the other conditions, their facilitations were scaled based on the results for transposition primes, which were also run under the yes/no procedure. To produce the scaled values, we compared the amount of facilitation for transposition primes under the go/no-go and the yes/no procedures, and multiplied the actual facilitations for the replacement primes by this factor, 1.7. Lower Panel: Mathematical model of data from Peressotti and Grainger (1995). The fit is based on a combination of fast parallel retinotopic priming, and slower serial abstract priming, as explained in the text

a reduced effect. For cross-position primes (312), facilitation is sensitive to exposure duration. Replacement primes (12d, 1d3, d23) induced a positional effect (12d produced significantly more facilitation than 1d3), while transposition primes (132, 321, 213) did not.

The fact that facilitation occurred for cross-position primes indicates that there are letter representations that can encode a given letter at different retinal locations and string positions. Indeed, based on duration effects, the authors concluded that there are two types of letter detectors, position-specific and position-independent. Position-specific detectors are activated first, followed by the position-independent detectors. Whitney (2001b) argued that the position-specific effect may not actually stem from string position, but rather may be retinotopic. Although primes and targets were presented in different font sizes (so letters in the same string position in the prime and target appeared at different retinal locations), there may still have been retinotopic overlap between letters at the same position, due to feature detectors with receptive fields large enough to span the differing retinal locations.

Indeed the data are well approximated under the following assumptions. The prime activates a retinotopic encoding in parallel (corresponding to the feature layer of the SERIOL model). The retinotopic encoding activates abstract letter units serially at the rate of about 15 ms a letter. For a 33 ms prime, only the first letter of the prime activates an abstract letter unit. For 50 ms, the first two letters of the prime activate abstract letter units, while all three abstract letter units are activated at 67 ms. The following values provide a reasonable fit to the data, shown in the lower panel of Figure 3. A retinotopic match (match on letter identity and position) provides 5 ms of facilitation per letter, and an abstract match provides an additional 5 ms of facilitation. (A retinotopic match does not necessarily cause an abstract match, because the number of abstract matches depends on exposure duration.) The total amount of priming is given by the sum across letters for retinotopic and abstract matches, up to a maximum facilitation of 25 ms.

For cross-position primes, the seriality explains the monotonic increase with exposure duration. The seriality also explains the difference in positional effects between transposition and replacement primes, for which presentation duration was 50 ms. Recall that only the first two positions of the prime activate abstract letter units at this duration. Replacement-prime 12d generates two abstract matches because the first two letters are in the target. However, 1d3 only produces one abstract match, because the letter in the second position is not present in the target. Similarly, d23 only generates one abstract match, because the first letter is not in the target. Thus, the number of abstract matches varies with replacement position, creating a positional effect. For transposition primes, the first two letters are present in the target in every condition (132, 321, or 213). Hence, all transposition primes provide one retinotopic and two abstract matches, giving no effect of the matched position.

This analysis leads to the prediction that the positional effect for replacement primes should disappear for a prime duration of 67 ms, because an abstract unit representing the prime's third letter should also become activated. That is, all replacement primes should yield two abstract matches at this longer duration.

Thus these data are consistent with SERIOL's proposal of sequential activation of abstract letter units. The other models cannot account for these data because they do not include abstract, location-independent letter units. Note that the phenomena cannot be explained by simply proposing that there are letter units with receptive fields spanning multiple string positions. Such units would lose important location/position information. There would have to be an additional mechanism for binding this information to such units, but the parallel models have not specified such a mechanism. (In the SERIOL model, string position is bound to the abstract letter units via firing order.) Of course, the parallel models do not rule out the possibility of abstract letter units, but such units do not play a direct role in the proposed orthographic processing.

PERCEPTUAL PATTERNS

Perceptual tasks provide another method for investigating orthographic encoding. In particular, it is of interest to note how perceptibility varies with string position, and to examine how the positional pattern interacts with alphanumeric status, visual field, and reading direction.

Two studies have examined perceptual patterns for letters versus non-alphanumeric characters in arrays of centrally presented stimuli (Hammond & Green, 1982; Mason, 1982), using a between-subjects design for the different stimulus types. Both studies found an external-character advantage for letters. That is, the first and last letters were processed more efficiently than the internal letters. Mason (1982) also showed an external-character advantage for number strings. However, both studies found that the advantage was absent for non-alphanumeric characters; the first and last characters were processed the *least* well, in line with their lower acuity. Thus symbols that do not normally occur in strings show a different perceptual pattern than alphanumeric characters. Therefore, the external-letter advantage cannot be due to a general reduction of lateral inhibition, because such a low-level effect should be present for any type of character. Any account of letter-string processing should then explain how the external-letter advantage arises.

For three-letter strings briefly presented to a single hemifield, the perceptual pattern varies with visual field. For left-to-right languages, the first letter is perceived much better than the third in the LVF, while the first and third letters are perceived equally well in the RVF (Hellige, Cohen, & Eng, 1995; Legge, Mansfield, & Chung, 2001). This asymmetry reverses with reading direction, as shown by studies in Hebrew (Eviatar, 1999) and Urdu (Adamson & Hellige, 2006). Thus these patterns are clearly contingent on the demands of encoding letter strings for reading, and therefore any model of orthographic processing should account for them.

As described in more detail elsewhere (Whitney, 2001a, 2007; Whitney & Cornelissen, 2005), the SERIOL model explains these patterns. The external-letter advantage arises as follows. An advantage for the initial letter comes from the directional inhibition at the (retinotopic) feature level, because the initial letter is the only letter that does not receive lateral inhibition. An advantage for the final letter arises at the (abstract) letter level, because the firing of the last letter is not terminated by a subsequent letter. This processing is specific to letter strings, explaining the lack of external-character advantage for non-alphanumeric characters.

However, if the initial-letter advantage arises at the feature level (i.e., prior to character identification), how could it be specific to letters? Recall that the subjects seeing non-alphanumeric characters saw only that type of stimulus, under the between-subjects design (Hammond & Green, 1982; Mason, 1982). Thus string-specific mechanisms would not have been activated in such subjects. This suggests that an initial-character advantage for non-alphanumeric arrays may emerge under a design in which alphabetic and non-alphanumeric stimuli are interspersed.

The LVF pattern of a strong advantage for the first over the third letter is explained by the proposed feature-level, left-to-right inhibition specific to the LVF/RH in left-to-right languages. In the absence of such inhibition, there is a shallower decrease in activation across the letters, yielding the more even RVF/LH pattern. For right-to-left languages, this directional inhibition should occur for the RVF/LH, explaining the reversal of trigram patterns with reading direction.

Whitney (2004b) proposed that these differing orthographic activation patterns are the cause of visual-field asymmetries in lexical decision. This theoretical account was later experimentally confirmed by using contrast manipulations to change orthographic activation patterns, resulting in reversal of normal asymmetries. Whitney and Lavidor (2004) demonstrated how to negate the length effect that is normally observed under LVF presentation (Young & Ellis, 1985). Whitney and Lavidor (2005) showed how to reverse the usual interaction between visual field and effect of orthographic-neighbourhood size (Lavidor & Ellis, 2002).

In the parallel models, activation levels at the letter layer either do not vary, or vary based on the offset from the centre of the corresponding receptive field. Thus they cannot explain the differing patterns for alphanumeric versus non-alphanumeric characters, nor the hemifield patterns.

Next we contrast the models at higher levels of processing. We examine arguments for and against serial lexical access, and then discuss the suitability of the proposed representations for lexical activation along both the lexical and sub-lexical routes.

SERIALITY

Because lexical-decision experiments have indicated that reaction times do not vary with string length (Frederiksen & Kroll, 1976; Richardson, 1976), it has generally been assumed that lexical access occurs in parallel. However, this reasoning is problematic. It depends on the assumption that the amount of time that it takes to reach response criterion after all the letters have fired (the *settling time*) is independent of the number of letters. However, if the greater amount of bottom-up input from a longer word were to yield a decreased settling time, this could cancel or override the increased time it takes for all the letters to fire. Thus there could be no length effect, or even a reversed length effect, under a serial encoding.

Length effects

Indeed, there are a number of studies consistent with the idea that longer strings may take more time to process at the orthographic level, but then take less time at the lexical level. In an EEG study, Hauk and Pulvermuller (2004) found that long words produced more activity than short words between 80–150 ms, while long words produced less activity than short words after 150 ms. Effects of word frequency first appeared around 150 ms, indicating that lexical access was in progress during this latter time frame. Thus the weaker activity for long words after 150 ms is consistent with reduced settling time. In MEG dipole source analyses, both Cornelissen, Tarkiainen, Helenius, and Salmelin (2003) (4-, 6- and 8-letter words) and Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin (1999) (single letters, 2-letter consonants and 4-letter words) showed a systematic decrease in latency with increasing string length for letter-string specific responses around $\sim$150ms after stimulus presentation. In another EEG study, Penolazzi, Hauk, & Pulvermuller (2007) found an interaction between length and frequency, and between length and semantic context at 120–180 ms; these interactions were not present at later intervals. This suggests that the initial phase of lexical activation overlapped with ongoing orthographic analysis, consistent with a serial letter encoding that yields incremental activation of lexical items.

Furthermore, a recent analysis (New, Ferrand, Pallier, & Brysbaert, 2006) of lexical-decision reaction times on a database of over 40,000 words (Balota et al., 2002) showed that RTs actually decrease as length increases from three to five letters; RTs are constant for words of five to seven letters, and RTs increase with length for seven or more letters. Hence, length has a facilitative, neutral, or inhibitory effect, depending on the range considered. It is highly unlikely that this variation reflects different methods of lexical access for words of different lengths. Also, Whitney and Lavidor (2004) showed that the length effect normally observed for LVF presentation can be negated by a suitable contrast manipulation. It is also highly unlikely that this manipulation changed the method of lexical access. Thus, the presence or absence of a length effect cannot be used to divine whether lexical access is serial or parallel.

Temporal manipulations

Rather, in order to test whether strings are processed serially, temporality should be manipulated directly. One such study has provided strong evidence for seriality. Nice and Harcum (1976) presented a six-letter pseudoword for 70 ms, which was replaced by another six-letter pseudoword for 30 ms. Subjects were to report the letters perceived. For the initial letter, subjects were more likely to report the letter from the first stimulus than the second, but for positions two through six, the pattern reversed; subjects were more likely to report the letter from the second stimulus than the first. Thus there was an interaction between position and temporal order, indicating serial processing. Note that this effect could not have been post-perceptual. If letter identification had proceeded in parallel, then there should have been no interaction between string position and order. Rather, the results indicate that different portions of the first and second strings were sampled at different times.

It is worth discussing in more detail whether the SERIOL model can account for the pattern of results in this study. The following data are taken from Figure 4 of Nice and Harcum (1976), in which a letter was counted as correct if it was reported within one position of its true position. For the first position, subjects reported the letter from the first stimulus in 25% of the trials, and the letter from the second stimulus in 17%. For the second position, reports from the first and second stimuli were 14% and 19%, respectively; in the third through sixth positions, reports from the first and second stimuli were approximately 10% and 17%. The letters were drawn from a pool of 12, so chance was about 8%. Thus report of the letters from the first stimulus fell to near chance in positions three to six.

For the first string position, the identification rate from the first stimulus was not near ceiling, while the rate from the second stimulus was above chance. Isn't this pattern incompatible with the idea that the first letter of the first stimulus was reliably processed before the second stimulus was presented? That is, if processing of the first letter was complete before the second stimulus even appeared, how could the first letter of the second stimulus influence report of the first letter? Why wasn't the recognition rate from the first stimulus higher?

This pattern is explained by the existing proposal that string processing occurs across multiple oscillatory cycles (Whitney & Lavidor, 2005). First we elaborate on the general proposal, and then discuss it in the context of the present experiment. In the following, we use the term 'scan' specifically to denote the conversion from a parallel representation to a serial representation as proposed in the SERIOL model; this is an automatic, subconscious process and should not to be confused with a scan driven by top-down processing.

Recall that serial letter firing is proposed to be driven by a theta cycle, which has a duration of about 200 ms. Thus, over a period of about 400 ms, the letters can be scanned twice. During the first theta cycle, the letters are activated based on bottom-up information, and lexical items are initially activated. During the next oscillatory cycle, the letters are scanned again in the presence of top-down lexical information, a single interpretation of the string is established, and the corresponding semantic representation is fully activated. This is consistent with evidence that lexical effects (e.g., frequency) are first visible in the EEG and MEG at about 150–180 ms, but semantic effects peak at 400–450 ms post-stimulus (i.e., the N400 component) (Penolazzi, Hauk, & Pulvermuller, 2007). That is, lexical effects first appear on the time scale of one theta cycle and peak after two theta cycles, consistent with two waves of activation.

Single cell recording indicates that a very brief stimulus (16 ms) results in elevated firing for about 300 ms (Rolls, Tovee, & Panzeri, 1999). Thus, although the stimuli in the present experiment were only exposed for a total of 100 ms, continued neural activity would have been sufficient to support multiple scans. The first scan would start across the first stimulus, allowing read out of the initial letters. Then the second stimulus would induce neural activity that was a combination of the two strings; the first scan would then operate over this combined representation for the remainder of the string. Then the second scan would operate on the combined representation for all positions, superseding the first scan. However, the second scan would be influenced by the first one, via top-down excitation from partially activated lexical items, and increased excitability of previously activated letters. Thus, there would be an increased probability of identifying the letters in the initial two positions as coming from the first string, but this tendency would not be absolute.

This serial account explains the pattern of experimental results. However, couldn't an activation gradient across a purely parallel representation explain the data? That is, perhaps activation level varies with position, and the way in which the two stimuli combine varies with activation level, so the relative contribution of each stimulus varies with position. While this possibility can't be completely ruled out, it would require the assumption of very particular non-linear interactions between the two stimuli in order to explain the present pattern. Identification of the first stimulus was relatively high for the first position, reduced for the second position, and flattened out to near chance for the remaining positions, while identification of the second stimulus was fairly flat across all positions. Thus the relative proportions of the two stimuli did not vary smoothly across positions; this non-linearity is inconsistent with a graded combinatory effect of activation levels. Rather, the pattern is predicted and most parsimoniously explained by the proposal that letters are identified serially.

Another experiment that manipulated temporality also provides support for serial processing. Adelman and Brown (2007) presented a four-letter word for variable durations (0 to 42 ms, in increments of 6 ms), followed by a mask and a two-alternative forced choice between the target and a foil word. The foil was of various types: (1) single-replacement: d234, 1d34, 13d4, or 123d; (2) double-replacement: dd34, 1dd4, 12dd, or d23d; (3) transposition: 2134, 1324, 1243, or 4231; or (4) all different: dddd. Thus the position of difference between the target and foil was parametrically varied in the foil types. Based on the interaction between accuracy, exposure duration, and position of difference, they found that latency of letter information increased with string position. Information about letter identity and order was available earlier for the initial letter than the internal letters, and earlier for the internal letters than the final letter. These results are consistent with a read-out of letter information from left to right on a millisecond time scale. It is interesting to note that there was no final-letter advantage. This is consistent with the SERIOL account of the final-letter advantage, which depends on the ability of the final letter unit to fire for a longer period of time than the internal letters. For very brief presentations, the final letter is not activated or does not receive enough input to sustain prolonged firing, yielding no advantage.

While the results of Adelman and Brown are consistent with serial processing, they could also be explained by a parallel mechanism with an activation gradient, in which the overall strength of the gradient grows with time. However, this would be inconsistent with the final-letter advantage observed in experiments with longer exposure durations (e.g., Hammond & Green, 1982; Mason, 1982), because the final letter would always be the *least* activated letter under a gradient account. Thus the Adelman and Brown results, when considered in combination with data from other experiments, are best explained via serial processing. Note that the alternative models under consideration cannot explain these results at all, as each is a parallel model without a monotonically decreasing activation gradient.

Oscillatory activity

In fact, the serial oscillatory mechanism was originally motivated by the need to explain the absence of a final-letter advantage at very brief exposures versus its presence at longer exposures, as discussed in Whitney and Berndt (1999). This proposal was based on theoretical considerations, not on experimental evidence for oscillatory activity. However, such evidence has subsequently emerged. Klimesch et al. (2001) showed that left occipital theta power increased with string-processing demands for normal, but not dyslexic, subjects. Krause et al. (2006) demonstrated that a theta power increase during lexical processing was specific to the visual modality of presentation. Bastiaansen, van der Linden, Ter Keurs, Dijkstra, and Hagoot (2005) found that left occipital theta power increased with string length.

The proposed coupling of gamma sub-cycles to a theta carrier wave predicts that gamma power should vary systematically with theta phase, because activity at the gamma frequency is taken to occur during a specific portion of the theta cycle. This coupling is taken to be a basic representational mechanism of the visual system (Whitney & Cornelissen, 2005), implying that gamma/theta co-variation should be present during both visual word and object recognition. In an EEG experiment, such co-variation has indeed been observed for visual object perception (Demiralp et al., 2007). To our knowledge, the gamma/theta relationship has not yet been investigated for string stimuli.

In sum, a survey of the experimental data shows no studies in conflict with seriality, and finds strong evidence for sequential processing. Furthermore, simulations presented in Whitney (2004a, 2007) show how serial activation of open-bigrams accounts for error patterns in aphasia (Berndt, Haendiges, & Mitchum, 2005; Whitney & Berndt, 1999), and for the presence and absence of positional effects in masked form priming, depending on the type of prime (Schoonbaert & Grainger, 2004; Grainger et al., 2006).

DIRECT LEXICAL ACCESS

Dehaene et al. (2005) argued that an open-bigram encoding could not provide a sufficiently robust encoding of letter order to support lexical recognition. They considered an on/off encoding with no edge units, arguing that such a representation fails to give a unique coding to strings with repeated letters. For example, the strings *sense*, *ensse* and *sensse* would all generate the same representation under these assumptions.

However, this argument does not apply to the open-bigram representation specified in Whitney (2004a), in which graded activations represent the separation of the component letters, edge units specify the external letters, and weights encode length information. To test the adequacy of this encoding, the string *sensse* was added to the most recent implementation of the bigram and word levels of the SERIOL model (Whitney, 2007). With no change to the parameter settings, the stimulus *sense* yielded **SENSE** as the winning word unit, while *sensse* activated **SENSSE** as the winner. Thus an open-bigram encoding can differentiate between similar strings with repeated letters, contrary to the claims of Dehaene et al. However, the difference in activation levels between **SENSSE** and **SENSE** was small. In the following section, we discuss evidence that orthographic encoding is actually more precise along the sub-lexical route.

Dehaene et al. (2005) solved the ambiguity problem by including units that encode higher-order sets of letters – bigrams activate four-grams, which activate words. However, such a representation means that there is zero similarity between similar words such as *care* and *core*. That is, the stimulus *core* would not activate the word-unit **CARE** at all, because it would not activate the four-gram **CARE**. Intuitively, it is obvious that *core* is more similar to *care* than to a word matching no letters, yet their proposed encoding does not capture this similarity. While it could be argued that *core* partially activates the four-gram **CARE**, such partial activations were not specified in their model. Once partial activations are considered, open-bigrams can attain a finer encoding of letter order (than under the on/off constraint), and there is no reason to include a higher-order encoding.

PHONOLOGY

Goswami and Ziegler (2006) argued that the demands of phonological processing should also be considered in constructing models of visual word recognition. They pointed out that parallel open-bigrams do not provide a suitable basis for learning grapheme-phoneme mappings, as open-bigrams are not phonologically relevant units.

However, their argument does not apply to the SERIOL model, because there is a serial, location-invariant representation of individual letters, which provides input to both the lexical and sub-lexical routes. The open-bigram representation is taken to be specific to the lexical route (Whitney & Cornelissen, 2005). Recall that the serial representation activates a grapho-phonological, syllabic representation along the sub-lexical route. Indeed, there is mounting evidence that the phonological representation is assembled serially (Carrerias, Ferrand, Grainger, & Perea, 2005; Roberts, Rastle, Coltheart, & Besner, 2003), which meshes nicely with the proposed serial letter encoding.

The proposed differences in orthographic encoding along the two routes suggest that letter order is encoded more reliably on the sub-lexical route. That is, open bigrams introduce ambiguity along the lexical route, while a syllabic representation provides a more veridical encoding of the string. Indeed, Frankish and Turner (2007) concluded that letter-order encoding is more precise along the sub-lexical route, based on studies involving Transposed-Letter (TL) nonwords formed by transposing two adjacent letters of a base word. They found that unpronounceable TL nonwords, such as *glvoe*, are more likely to be misperceived as actual words than are pronounceable TL nonwords, such as *golve*. This result is quite striking, as it contradicts the usual finding that orthographically illegal nonwords are easier to reject than pronounceable nonwords.

Based on a series of experiments, Frankish and Turner (2007) concluded that the sub-lexical representation generated by a pronounceable TL nonword inhibits the mismatching base word (which was activated via the lexical route); for unpronounceable TL nonwords, there is no sub-lexical representation to inhibit the base word, so more lexicalisation errors occur. (The pattern of experimental results supported this account better than the alternative possibility that unpronounceable TL nonwords directly cause more activation of the base word.) Thus the sub-lexical route generates distinct representations for *glove*, *glvoe*, and *golve*, implying a precise encoding of letter order, while the lexical route has more difficulty distinguishing between these stimuli.

Despite the suitability of SERIOL's representations for the demands of phonological processing, Goswami and Ziegler (2006) argued that 'This solution ignores data showing that phonology affects the lexical route, such as body-neighborhood effects in lexical decision (Ziegler & Perry, 1998)'. It is unclear what is meant by this statement. First, Whitney (2004b) specifically discussed the data presented by Ziegler and Perry (1998), explaining in detail how the SERIOL model explains their findings. Second, the general issue of interaction between the lexical and sub-lexical routes is orthogonal to the question of how letter order is encoded. Presumably, the lexical and sub-lexical routes converge onto the same lexical representations. The usual assumption of top-down connections from the lexical representations back to letters and phonemes would cause interaction between the routes, because orthographically driven lexical information could then affect the phonemic representation, and vice versa. This connectivity pattern is independent of how letter position is encoded.

CONCLUSION

We have seen a range of relevant data that the parallel models are unable to explain. Because these models do not specify a level of abstract letter units, they cannot account for letter-priming results and do not provide a representation suitable for phonological processing. Because the alternative models do not include an activation gradient or serial processing, they are unable to explain observed patterns of letter perceptibility, the influence of temporal manipulations, the effect of contrast manipulations on visual-field asymmetries, and the existence of oscillatory phenomena in string processing. In contrast, the SERIOL model predicts and explains all of these findings, while specifying representations that are suitable for both lexical and phonological processing. For a discussion of the resulting implications for dyslexia and visual object recognition, see Whitney and Cornelissen (2005).

REFERENCES

Adamson, M. M., & Hellige, J. B (2006). Hemispheric differences for identification of words and nonwords in Urdu-English bilinguals. *Neuropsychology*, *20*, 232–248.

Adelman, J. S. & Brown, G. D. A. (2007). The time course of letter position and identity information in visual word identification. Submitted.

Balota, D. A., Cortese, M. J., Hutchison, K. A., Neely, J. H., Nelson, D., Simpson, G. B., & Treiman, R. (2002). The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords. St. Louis: Washington University. Available at elexicon.wustl.edu.

Bastiaansen, M.C., van der Linden, M., Ter Keurs, M., Dijkstra, T., & Hagoot, P. (2005). Theta responses are involved in lexical-semantic retrieval during language processing. *Journal of Cognitive Neuroscience*, *17*, 530–541.

Berndt, R. S., Haendiges, A. N., & Mitchum, C. C. (2005). Orthographic effects in the word substitutions of aphasic patients: an epidemic of right neglect dyslexia? *Brain and Language*, *9*, 55–63.

Buchholz, J., & Aimola Davies, A. (2006). Do visual attentional factors contribute to phonological ability? Studies in adult dyslexia. *Neurocase*, *12*, 111–121.

Carreiras, M., Ferrand, L., Grainger, J., & Perea, M. (2005). Sequential effects of phonological priming in visual word recognition. *Psychological Science*, *16*, 585–589.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, *91*, 77–111.

Cornelissen, P. L., Richardson, A. J., Mason, A. J., & Stein, J. F. (1995). Contrast sensitivity and coherent motion detection measured at photopic luminance levels in dyslexics and controls. *Vision Research*, *35*, 1483–1494.

Cornelissen, P. L., Tarkiainen, A., Helenius, P., & Salmelin, R. (2003). Cortical effects of shifting letter-position in letter-strings of varying length. *Journal of Cognitive Neuroscience*, *15*, 731–748.

Davis, C. (1999). *The Self-Organizing Lexical Acquisition and Recognition (SOLAR) model of visual word recognition*. Doctoral Dissertation. University of New South Wales.

Davis, C. J., & Bowers, S. J. (2006). Contrasting five different theories of letter-position coding: Evidence from orthographic similarity effects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 535–557.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Science*, *9*, 335–341.

Demiralp, T., Bayraktaroglu, Z., Lenz, D., Junge, S., Busch, N. A., Maess, B., Ergen, M., & Herrmann, C. S. (2007). Gamma amplitudes are coupled to theta phase in human EEG during visual perception. *International Journal of Psychophysiology*, *64*, 24–30.

Eviatar, Z. (1999). Cross-language tests of hemispheric strategies in reading nonwords. *Neuropsychology*, *13*, 498–515.

Facoetti, A., & Molteni, M. (2001). The gradient of visual attention in developmental dyslexia. *Neuropsychologia*, *39*, 352–357.

Frankish, C., & Turner, E. (2007). SIHGT and SUNOD: The role of orthography and phonology in the perception of transposed letter anagrams. *Journal of Memory and Language*, *56*, 189–211.

Frederiksen, J., & Kroll, J. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 361–379.

Goswami, U., & Ziegler, J. C. (2006). A developmental perspective on the neural code for written words. *Trends in Cognitive Sciences*, *10*, 142–143.

Grainger, J., Granier, J. P., Farioli, F., Van Assche, E., & van Heuven, W. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 865–884.

Grainger, J., & van Heuven, W. (2003). Modeling letter position encoding in printed word perception. In P. Bonin (Ed.), *The mental lexicon* (pp. 1–24). New York: Nova Science Publishers.

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a whole? *Trends in Cognitive Sciences*, *8*, 58–59.

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

Hauk, O., & Pulvermuller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, *115*, 1090–1103.

Hellige, J. B., Cowen, E. L., & Eng, T. L. (1995). Recognition of CVC syllables from LVF, RVF, and central locations: Hemispheric differences and interhemispheric interactions. *Journal of Cognitive Neuroscience*, *7*, 258–266.

Klimesch, W., Doppelmayr, M., Wimmer, H., Gruber, W., Rohm, D., Schwaiger, J., & Hutzler, F. (2001). Theta band power changes in normal and dyslexic children. *Clinical Neurophysiology*, *112*, 1174–1185.

Krause, C. M., Gronholm, P., Leinonen, A., Laine, M., Sakkinen, A. L., & Soderholm, C. (2006). Modality matters: The effects of stimulus modality on the 4- to 30- Hz brain electric oscillations during a lexical decision task. *Brain research*, *1110*, 182–192.

Lavidor, M., & Ellis, A. (2002). Word length and orthographic neighborhood size effects in the left and right cerebral hemispheres. *Brain and Language*, *80*, 45–62.

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.

Lisman, J. E., & Idiart, M. A. P. (1995). Storage of 7 ± 2 short-term memories in oscillatory subcycles. *Science*, *267*, 1512–1515.

Livingstone, M. S., Rosen, G. D., Drislane, F. W., & Galaburda, A. M. (1991). Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proceedings of the National Academy of Sciences*, *88*, 7943–7947.

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

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Re-examining word length effects in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin and Review*, *13*, 45–52.

Nice, D. S., & Harcum, E. R. (1976). Evidence from mutual masking for serial processing of tachistoscopic letter patterns. *Perceptual and Motor Skills*, *42*, 991–1003.

Penolazzi, B., Hauk, O., & Pulvermuller, F. (2007). Early semantic context integration and lexical access as revealed by event-related brain potentials. *Biological Psychology*, *74*, 374–388.

Peressotti, F., & Grainger, J. (1995). Letter position coding in random consonant arrays. *Perception and Psychophysics*, *57*, 875–890.

Richardson, J. (1976). The effects of stimulus attributes on latency of word recognition. *British Journal of Psychology*, *67*, 315–325.

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 and Review*, *10*, 405–414.

Rolls, E. T., Tovee, M. J., & Panzeri, S. (1999). The neurophysiology of backward visual masking: Information analysis. *Journal of Cognitive Neuroscience*, *11*, 300–311.

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.

Sperling, A. J., Zhong-lin, L., Manis, F. R., & Seidenberg, M. S. (2003). Selective magnocellular deficits in dyslexia: A "phantom contour" study. *Neuropsychologia*, *41*, 1422–1429.

Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter-string perception in the human occipito-temporal cortex. *Brain*, *122*, 2119–2131.

Whitney, C. (2001a). 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. (2001b). Position-specific effects within the SERIOL framework of letter-position coding. *Connection Science*, *13*, 235–255.

Whitney, C. (2004a). *Investigations into the neural basis of structured representations*. Doctoral Dissertation. University of Maryland.

Whitney, C. (2004b). Hemisphere-specific effects in word recognition do not require hemisphere-specific modes of access. *Brain and Language*, *88*, 279–293.

Whitney, C. (in press). A comparison of the SERIOL and SOLAR theories of letter-position encoding. *Brain and Langurage*.

Whitney, C. (2007). *Supporting the serial in the SERIOL model*. Manuscript submitted for publication.

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*, *51*, 179–213.

Young, A. W., & Ellis, A. W. (1985). Different methods of lexical access for words presented in the left and right visual hemifields. *Brain and Language*, *24*, 326–358.

Ziegler, J. C., & Perry, C. (1998). No more problems in Coltheart's neighborhood: Resolving neighborhood conflicts in the lexical decision task. *Cognition*, *68*, B53–62.