The SERIOL2 Model of Orthographic Processing

Carol Whitney and Yuval Marton

Self-Published Manuscript

6/7/13

The SERIOL2 Model of Orthographic Processing

Carol Whitney and Yuval Marton

Abstract

The SERIOL model of orthographic analysis proposed mechanisms for converting visual input into a serial encoding of letter order, which involved hemisphere-specific processing at the retinotopic level. As a test of SERIOL predictions, we conducted a consonant trigram-identification experiment, where the trigrams were briefly presented at various retinal locations. The accuracy data exactly matched the SERIOL predictions. To further test the SERIOL model, we conducted the trigram experiment in the Hebrew alphabet (read from right to left) with native Hebrew speakers. However, the SERIOL predictions were not fully confirmed for Hebrew. Therefore we revised the SERIOL model, resulting in the SERIOL2 model presented here. SERIOL2 re-specifies how the retinotopic representation is transformed into a serial encoding of letter order. We present spiking-neuron simulations to illustrate how SERIOL2 accounts for the trigram data.

Keywords: Visual Word Recognition, Orthographic Processing, Letter Perception, Spiking Neuron Model

1.0 Introduction

Recent years have witnessed an explosion of interest in the question of how letter position is encoded in orthographic processing during visual word recognition, as investigated in a range of computational, behavioral and imaging studies (e.g., Adelman, 2011; Binder, Medler, Westbury, Liebenthal, & Buchanan, 2006; Carreiras, Duñabeitia, & Molinaro, 2009; Davis & Bowers, 2006; Davis, 2010; Dehaene, Cohen, Sigman, & Vinckier, 2005; Frankish & Barnes, 2008; Frost, 2012; Gomez, Ratcliff, & Perea, 2008; Grainger, Granier, Farioli, Van Assche, & van Heuven, 2006; Holcomb & Grainger, 2006; Perea & Lupker, 2003; Velan & Frost, 2009; Whitney & Cornelissen, 2005; Whitney, 2001).

It is of critical importance to gain a complete and accurate model of orthographic processing, as a growing number of studies indicate that developmental dyslexia is associated with abnormal orthographic representations (e.g., Bosse, Tainturier, & Valdois, 2007; Helenius, Tarkiainen, Cornelissen, Hansen, & Salmelin, 1999; Van den Broeck & Geudens, 2012; van der Mark et al., 2009). We must first understand skilled orthographic processing in exquisite detail before we can hope to characterize the nature of developmental deficiencies.

This article focuses on the lower levels of orthographic processing, addressing the question of how a retinotopic representation is converted into an abstract (location-invariant) encoding of letter order. To motivate the issues addressed herein, we first briefly review the neural architecture of the reading network.

1.1 Overview of the Reading Network

Imaging studies have identified brain areas involved in orthographic processing, and have provided information about the nature of representations in some areas, as summarized in Figure 1. Two major processing pathways for visual word recognition are simultaneously active when individuals process printed words (Fiebach, Friederici, Müller, & von Cramon, 2002; S. M. Wilson et al., 2009; T. W. Wilson et al., 2007). On the ventral occipitotemporal pathway, an orthographic representation is thought to activate a lexicosemantic encoding, which activates a phonological representation. On the dorsal occipito-parieto-frontal pathway, an orthographic representation is thought to directly activate a phonological representation.

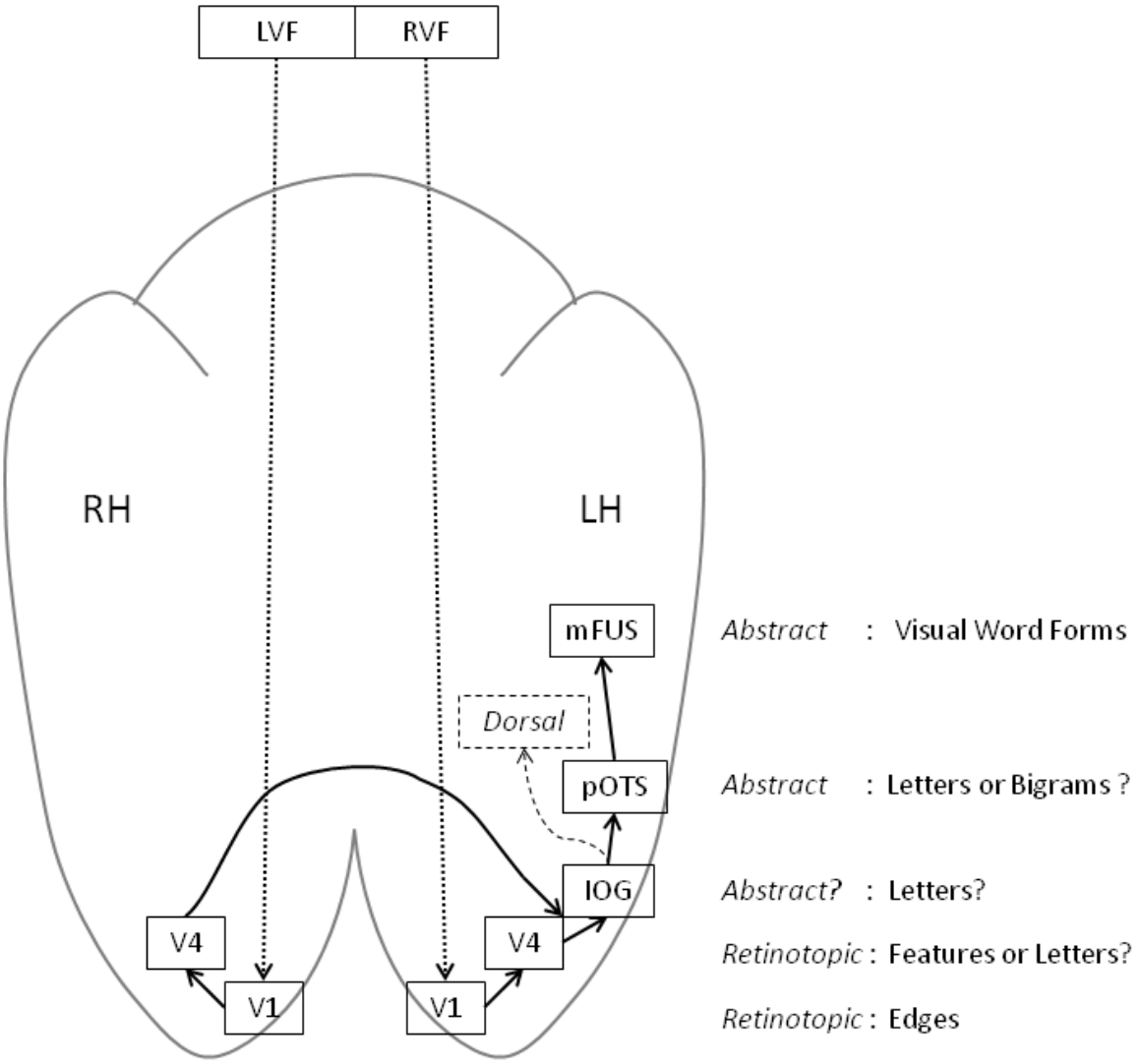

Figure 1: Organization of the reading network as viewed from the ventral surface of the brain. See text for abbreviations. Note that the left hemisphere is on the right. Cortical areas are shown in their approximate x, y locations. V1 projects to V4 via areas V2/V3, which are not shown due to space limitations. Notations on the right side of the figure describe known and/or suspected (tagged by '?') aspects of the orthographic representation in each area. Beyond the IOG, the two reading pathways diverge. The ventral pathway projects to the pOTS and mFUS. The dorsal pathway, which supports phonological processing, projects to temporo-parietal areas.

The cortical analysis of a letter string begins in visual area V1. ERP studies have established that the representation of the fovea in V1 is split across the hemispheres (Jordan, Fuggetta, Paterson, Kurtev, & Xu, 2011; Martin, Thierry, Démonet, Roberts, & Nazir, 2007). That is, stimuli just to the left of fixation project to V1 in the right hemisphere (RH), and stimuli to the right of fixation project to V1 in the left hemisphere (LH). Studies in which letter strings are presented unilaterally to the left visual field (LVF) or right visual field (RVF) suggest that orthographic processing remains lateralized to the contralateral hemisphere up to visual area V4 (Ben-Shachar, Dougherty, Deutsch, & Wandell, 2007; Cohen et al., 2000). It is well known that visual areas V1-V4 are retinotopically organized, and that V1 encodes oriented edges while V4 encodes complex combinations of edges. It is not known whether V4 encodes whole letters or sub-letter features during orthographic processing.

An MEG study (Barca et al., 2011) employing unilateral orthographic stimuli and "virtual electrodes" centered on bilateral Middle Occipital Gyrus (MOG) found strong activation of left MOG for both LVF and RVF stimuli, and strong activation of right MOG only for LVF stimuli. Therefore, the authors concluded that LH and RH orthographic representations converge near left MOG. As discussed next, recent studies have identified left Inferior Occipital Gyrus as a key area of orthographic processing. The MOG "virtual electrode" likely included signals from the IOG, and we assume that the MEG results reflect convergence near left IOG.

During visual word recognition, the left IOG is functionally connected to the superior temporal gyrus and inferior parietal lobe on the dorsal phonological pathway, and to occipitotemporal areas on the ventral pathway (Richardson, Seghier, Leff, Thomas, & Price, 2011; Seghier et al., 2012). This pattern suggests that the two reading pathways diverge after left IOG. Therefore, the IOG likely provides an orthographic representation that is suitable for phonological analysis on the dorsal pathway and whole-word recognition on the ventral pathway. An abstract encoding of individual letters would meet these requirements, for example. Indeed, an fMRI multivariate-pattern study showed that left (but not right) IOG encodes pseudoword identity independent of font, consistent with an abstract letter encoding (Nestor, Behrmann, & Plaut, 2012).

Beyond the IOG, we focus on the ventral pathway, where the left posterior Occipitotemporal Sulcus (pOTS) and mid Fusiform gyrus (mFUS) have been identified as important orthographic regions (Ben-Shachar et al., 2007; Cohen & Dehaene, 2004; Cohen et al., 2000; Mano et al.,

2012; Thesen et al., 2012). Masked-priming fMRI studies have provided information about representations in these areas. The pOTS region houses a representation that is not sensitive to retinal location, letter case, or word identity, indicating an abstract prelexical representation (Dehaene et al., 2004). The mFUS houses a representation that is sensitive to word identity, suggesting that it encodes Visual Word Forms (VWFs) (Glezer, Jiang, & Riesenhuber, 2009). Furthermore, the pOTS region is sensitive to the number of letters in a string and does not differentiate between consonant strings and words, while the mFUS region is not sensitive to string length and shows a stronger response to words than consonant strings (Thesen et al., 2012). These results also indicate prelexical and lexical representations in the pOTS and mFUS, respectively.

Although a recent study has shown sensitivity to retinal location of orthographic stimuli in the mFUS region (Rauschecker, Bowen, Parvizi, & Wandell, 2012), the range of retinal locations investigated in that study were well outside those normally used in reading. The present research is concerned with the formation of a location-invariant representation from retinal locations normally utilized in skilled reading; we assume that pOTS and mFUS support abstract representations of orthographic stimuli located on the horizontal meridian near fixation.

To summarize, retinotopic processing of contralateral orthographic stimuli occurs in visual areas V1-V4. The bilateral retinotopic areas likely converge near left IOG, which provides orthographic input to the dorsal and ventral reading pathways. Along the ventral pathway, the pOTS encodes a prelexical orthographic representation, which activates VWFs in the mFUS.

Much of the debate surrounding models of orthographic processing has focused on the nature of the highest-level prelexical representation, which would correspond to the encoding in left pOTS. The SERIOL (Sequential Encoding Regulated by Inputs to Oscillatory Letter-units) model was the first to propose an encoding based on not-necessarily-contiguous, ordered letter pairs (Whitney & Berndt, 1999; Whitney, 2001), which have come to be known as open-bigrams (Grainger & Whitney, 2004). For example, the stimulus "bird" would activate open-bigrams *BI, IR, RD, BR, ID, BD*. This proposal was originally driven by masked-priming studies and aphasics' error patterns, which indicated that letter-position encoding is sensitive to relative order, and is not position-specific (Whitney & Berndt, 1999; Whitney, 2001). Additional researchers have adopted this idea (Dehaene et al., 2005; Grainger et al., 2006), including open-

bigrams in their models. In contrast, other models propose that the highest-level prelexical encoding is based on a representation of individual letters (Davis, 2010; Gomez et al., 2008; Norris & Kinoshita, 2012).

A full model of orthographic encoding should specify not only the type of representation activating VWFs, but also how this representation is formed from the lower levels of visual input. The SERIOL model (Whitney & Cornelissen, 2005; Whitney, 2001) also addressed this issue, specifying how the early retinotopic encoding of a letter string is converted into the open-bigram representation. In brief, the SERIOL model proposes the following: (1) A serial encoding of letter order (in left IOG) activates open-bigrams (in left pOTS). (2) The serial encoding is induced by a parallel activation gradient at the retinotopic feature level (in bilateral V4). (3) The formation of this activation gradient requires hemisphere-specific patterns of excitation (from V1 to V4) and lateral inhibition (within V4).

1.2 Overview of the Present Study

This article focuses on the implications of assumptions (2) and (3) above. As discussed in more detail in Section 2.1, SERIOL specified that, for a language read from left to right, RH letter-feature representations inhibit letter-features at retinal locations to the right, but LH letter-feature representations do not provide such unidirectional lateral inhibition. Hence a letter presented to the LVF/RH should be strongly inhibited by a letter to the left, but not the right. In contrast, a letter presented RVF/LH should not be strongly inhibited by a letter to the left or right.

Exactly such a pattern was observed in a recent study of letter perception (Grainger, Tydgat, & Isselé, 2010), although the authors did not acknowledge that these results were predicted by the SERIOL model. To investigate SERIOL predictions in more detail, we performed a perceptual experiment in which consonant trigrams were briefly presented (for ~67 ms) at various retinal locations. This trigram experiment was conducted in the Latin alphabet with native English speakers. For convenience, we refer to these stimuli as “English” trigrams. The English results precisely matched the SERIOL predictions.

To further test the SERIOL model, we then conducted the trigram experiment in the Hebrew alphabet (a script read from right to left) with native Hebrew speakers. SERIOL predicted that the Hebrew accuracy pattern should be a mirror image of the English pattern. However, this

prediction was only partially confirmed. Furthermore, the Hebrew data yielded a very surprising result: accuracy for the initial letter of a trigram was considerably *lower* when this letter was directly fixated than when it was located five letter-widths to the right of fixation!

Based on the implications of these unexpected findings, we modify the SERIOL model to account for both the English and Hebrew results. Originally, SERIOL stood for Serial Letter Encoding Regulated by Inputs to Oscillatory Letter-units. As we will see, the modified model is a direct descendant of SERIOL, but the *Inputs to Oscillatory* aspect of SERIOL is no longer applicable. Therefore the new model is dubbed SERIOL2, under the provision that the acronym now stands for SERIalization Of Letters. SERIOL2 re-specifies how the retinotopic encoding is converted into a serial encoding of letter order; beyond the letter level, SERIOL2 remains essentially the same as SERIOL. Via a simulated spiking-neuron network, we show that SERIOL2 explains the observed trigram patterns for both reading directions.

The organization of this paper is as follows. First we review the SERIOL model. Then we present the English and Hebrew trigram experiments. Next we consider the implications of the trigram experiments and formulate the SERIOL2 model of skilled orthographic processing. We then introduce a neural network that implements key aspects of SERIOL2, and present simulations of normal string processing and the trigram experiments. In the General Discussion, we consider implications of SERIOL2 for VF asymmetries in lexical-level processing, consider related models, and outline directions for future research.

2.0 Review the SERIOL Model

In describing neural models, it can become confusing as to whether one is referring to a stimulus, or the neural encoding of the stimulus. For clarity in the following, we use quotation marks to denote a stimulus, capitalized words (such as Letters or Features) to indicate a category of neural representation, and italics to denote the neural representation of a particular stimulus. For example, the stimulus "BIRD" consists of letters 'B', 'I', 'R', and 'D', and this stimulus should activate Letters *B, I, R,* and *D* in the brain.

As discussed in the Introduction, the early cortical representation of a letter string is retinotopic. For example, consider the stimulus "CAT". Fixation on the 'C' would yield a completely different pattern of retinotopic activity than fixation on the 'T'. Yet both of these activity patterns

should activate the VWF *CAT* on the ventral pathway, and the phonemic encoding */k a t/* on the dorsal pathway. A key goal of models of orthographic processing should be to specify how the transformation from a retinotopic to a location-invariant (abstract) encoding occurs. The lower levels of the SERIOL model address this issue.

2.1 SERIOL Model

The model is comprised of Edge, Feature, Letter, Open-Bigram, and Word areas. The term *activation level* will be used to denote the total amount of neural activity devoted to representing a letter stimulus within a given area. Activation level increases with the number of active neurons and their firing rate. For example, given the stimulus "ON", the activation level of O in the Feature area corresponds to the summed activity of the Features driven by the 'O'.

Edges

The Edge area models bilateral V1/V2. The model highlights three known attributes of these cortical areas: (1) Lower visual areas are retinotopically organized. (2) The representation of the left and right visual hemifields is initially split across the hemispheres, as discussed in the Introduction. (3) The retinotopic encoding is subject to an *acuity gradient;* that is, the amount of cortical tissue representing a visual stimulus of a given size decreases as the eccentricity of the stimulus increases. The acuity gradient implies that letter activation level decreases as the letter's distance from fixation increases, because fewer neurons are activated by stimuli farther from fixation.

Letter Features

The Feature area models bilateral V4. V4 is also retinotopic, divided into RH and LH representations, and subject to the acuity gradient. However, the SERIOL model proposes that reading acquisition causes hemisphere-specific processing of letter Features. This processing converts the acuity gradient of the Edge into a monotonically-decreasing activation gradient in the Feature area, dubbed the locational gradient. The locational gradient is an activation pattern wherein activation level is highest for the Features of the first letter and decreases across the Feature representation of the string, independent of the retinal location of the string. This is accomplished by increasing or decreasing the firing rate of relevant cells.

SERIOL assumes that this processing is learned during reading acquisition. The locational gradient is initially imposed by a top-down attention gradient, and the visual system then learns to create this activation pattern in a bottom-up manner. SERIOL specifies the bottom-up processing necessary to convert the acuity gradient into the locational gradient. This processing consists of three transformations. Figure 2 presents a network architecture that implements the transformations, while Figure 3 illustrates the effects of these transformations on activation patterns. These Figures and the following description of the transformations are for a language read from left to right. For a language read from right to left, the processing is reversed across the hemispheres.

The first transformation is stronger Edge → Feature (bottom-up) excitation in the LVF/RH than the RVF/LH. This raises the activation level of Features encoding all LVF letters. In particular, it brings the Features of the first letter to a high activation level.

The second transformation is strong left-to-right lateral inhibition within the LVF/RH Feature area. This is required to invert the acuity gradient, which *increases* from left to right, into a gradient that *decreases* from left to right. That is, the Features of each letter inhibit the Features of letters falling to the right. For example in Figure 3, the Features of the first letter inhibit the Features of the second letter, while the Features of the first and second letters inhibit the Features of the third letter. As a result, activation level in the RH decreases toward fixation. Such unidirectional lateral inhibition is not necessary within the RVF/LH, because the acuity gradient decreases from left to right.

The third transformation is cross-hemispheric inhibition within the Feature area. All LVF/RH Features inhibit all RVF/LH Features. This brings the activation level of all RVF/LH Features lower than all LVF/RH Features, “joining” the two hemispheric gradients into a monotonically-decreasing gradient.

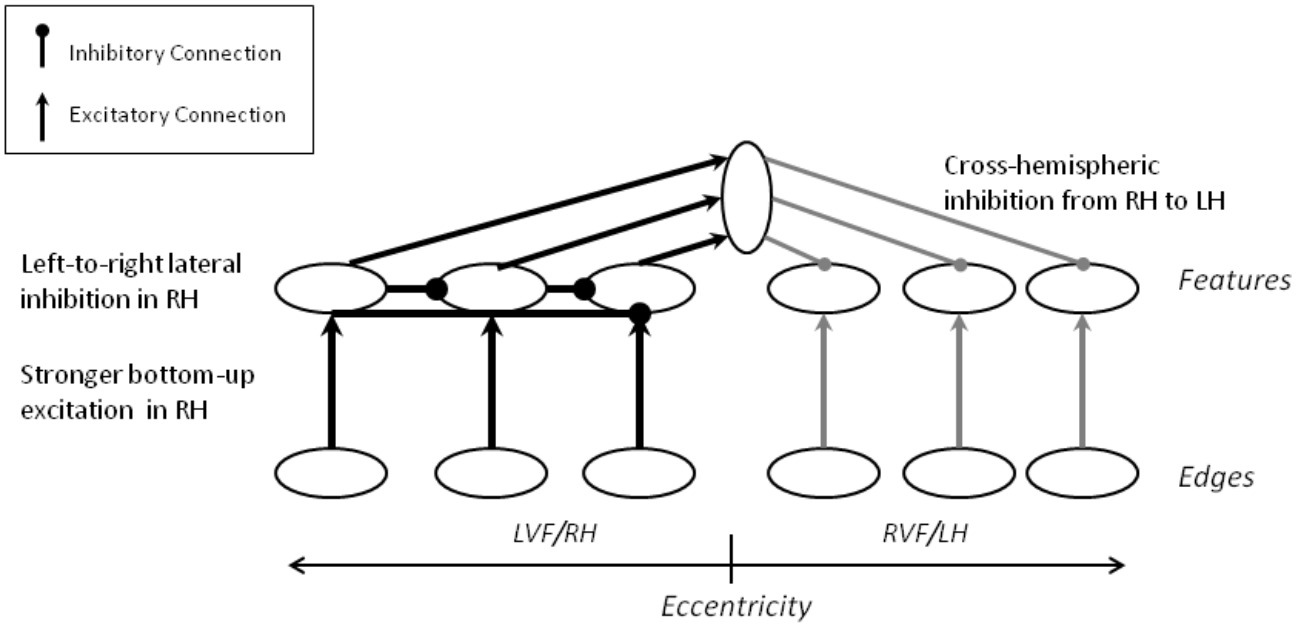

Figure 2: Architecture of a network that converts the acuity gradient (of the Edge area) into the locational gradient (of the Feature area) for a left-to-right language. Darker lines represent stronger connections. Each horizontal oval represents the set of edges or features comprising a letter. The vertical oval represents a pool of inhibitory interneurons.

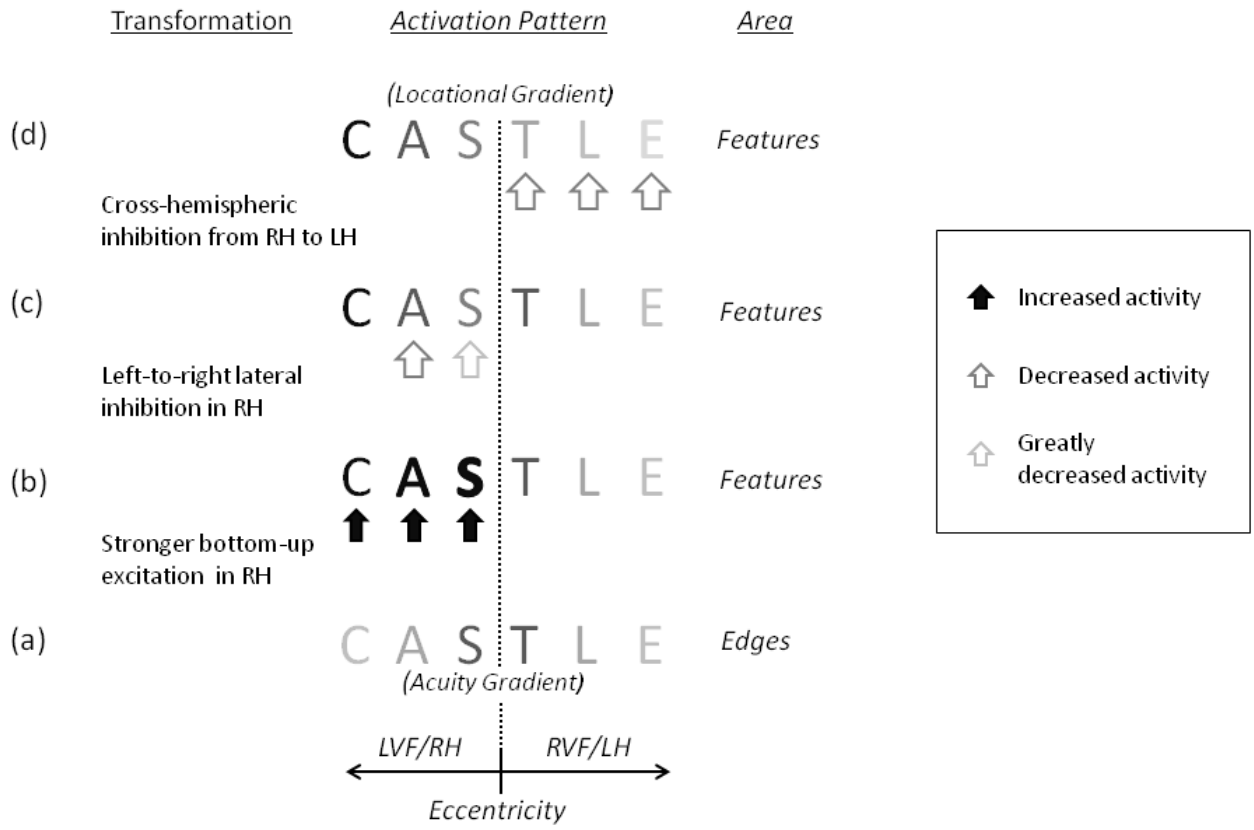

Figure 3: Illustration of the three transformations of locational-gradient formation, for the centrally-fixated stimulus "CASTLE". Darker and wider letters represent higher activation levels. Rows (b-d) illustrate the effect of each transformation on the activity pattern from the row below, where an arrow highlights a change in activation level. Note that the Feature area is repeated in order to present the effect of each transformation separately; this is not meant to imply multiple Feature areas. The transformations are shown sequentially for illustrative purposes; they would actually occur interactively.

Row (a) illustrates the activation pattern in the Edge area due to the acuity gradient, wherein activation level decreases with increasing distance from fixation. Row (b) illustrates the effect of stronger Edge→Feature excitation in the RH than LH, which is necessary for the first letter to attain a high activation level. Row (c) illustrates the effect of unidirectional lateral inhibition within the RH Feature area. The second letter is moderately inhibited (by the first letter), while the third letter is strongly inhibited (by the first and second letters). As result, activation level now decreases from left to right within the RH. In the LH, activation level already decreases from left to right (due to the acuity gradient), so unidirectional inhibition is not necessary. Row (d) illustrates the effect of cross-hemispheric inhibition. All RH letters inhibit all LH letters, such that the activation level of the fourth letter becomes lower than the third letter. The result is the monotonically-decreasing locational gradient.

Note that these Features are taken to be specific to letter processing. We assume that multiple Feature-Detector neurons respond to the same feature in a stimulus. Some of these Feature Detectors are connected in the way proposed by the SERIOL model (the present Features), to support the specializations required for orthographic processing. Other Feature Detectors are connected in a manner that supports general object recognition, and are not subject to locational-gradient formation (General Features). For example, the symbol string "#%&" would activate both Features and General Features. The Features would fail to activate Letters, while the General Features would activate the corresponding Symbols. Hence, the locational gradient does not apply to symbol strings (or to any other non-letter objects, except possibly numbers).

Letters

The Letter area corresponds to left IOG, and is comprised of abstract (non-retinotopic) letter representations. Letters fire sequentially. That is, the Letter encoding the first letter fires, then the Letter encoding the second letter fires, etc. The induction of this serial encoding is based on the proposal of a general brain mechanism in which item order is encoded in successive gamma cycles (60 Hz) of a theta cycle (5 Hz) (Lisman & Idiart, 1995); SERIOL adapts this mechanism to encode letter order, wherein successive Letters fire ~16 ms apart (i.e., one gamma cycle apart).

Specifically, this firing pattern is induced by the interaction of the locational gradient with synchronous sub-threshold theta oscillations of the Letters' membrane potential. It is assumed that the theta oscillation is reset by a saccade or stimulus onset, such that the excitability of Letters is lowest when input from the Feature area first reaches the Letter area. The Letter corresponding to the first letter receives the most input from the Feature area (due to the locational gradient); this Letter crosses threshold and fires first. As the excitability of the Letters increases over time (due to the theta oscillation), the Letter receiving the next most input (i.e., the Letter encoding the second letter) crosses threshold and fires next, etc. See Whitney and Berndt (1999) for details and simulations.

This temporal encoding of letter order is a location-invariant representation. Therefore, it provides suitable input for both the dorsal and ventral pathways. Accordingly, processing in the SERIOL model bifurcates following the Letter area. On the dorsal pathway, the sequence of

Letters is parsed into a graphemic encoding, which is mapped to a phonemic encoding. We focus on the ventral pathway, where Letters activate Open-Bigrams.

Open-Bigrams

The Open-Bigram area corresponds to left pOTS. Open-Bigram *XY* is activated if *X* fires and then *Y* fires within ~50 ms. The activation level of an Open-Bigram decreases as the interval between between the firing of the constituent Letters increases. In the "bird" example, *BI* would attain a higher activation level than *BR*. SERIOL also assumes Edge Bigrams. For example, "bird" would activate Edge Bigrams **B* and *D**.

Words

The Word area corresponds to mFUS, and encodes VWFs. The Open-Bigrams activated by a given word have excitatory connections with the corresponding VWF. For example, **B*, *BI, BR, BD, IR, ID, RD* and *D** have excitatory connections to *BIRD;* other Open-Bigrams and Edge Bigrams either have no connections or inhibitory connections to *BIRD*. The SERIOL model is summarized in Figure 4.

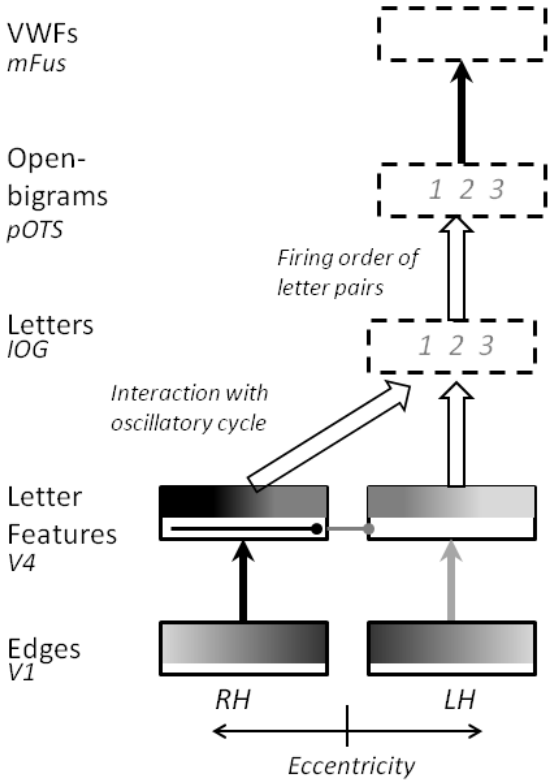

Figure 4: Diagram of the SERIOL model. Boxes with solid edges represent retinotopic representations, while boxes with dashed edges represent abstract representations. Darker grey levels represent stronger activation, excitation, or inhibition. Solid arrows represent excitatory connections, while open arrows represent more complex feed-forward processing. Horizontal capped lines represent lateral inhibition. The grey-level gradients indicate activation level across retinal location. The V1 gradient corresponds to the acuity gradient, while the V4 gradient corresponds to the locational gradient. The abstract serial letter representation in IOG also projects to the dorsal pathway (not shown), where it is parsed into graphemes and mapped to phonemes.

3.0 Trigram Experiment with Native English Speakers

Next we consider an experimental test of the proposed processing at the Feature and Letter levels in the SERIOL model. This experiment utilized trigram identification, wherein a consonant trigram is briefly presented and then masked, and the subject is to report all of the letters. A similar task has previously been used to measure the "visual span" for reading (Dubois, De Micheaux, Noël, & Valdois, 2007; Kwon, Legge, & Dubbels, 2007; Legge, Mansfield, & Chung, 2001). However, our intent is different. We are interested in measuring the effect of a letter's within-string position on perceptual accuracy, across retinal locations. In particular, we wish to see how the effect of position varies across visual fields, in order to evaluate the predictions of the SERIOL model. (Note that the term *location* will always refer to a letter's retinal location with respect to fixation, while *position* will refer to a letter's position with respect to the string.)

Because we are interested in the effects of within-string position, we utilize a post-mask that extends beyond the stimulus, in order minimize any general advantage for edge letters. Because we are interested in retinotopic processing normally utilized in reading, we limit our analysis to locations near fixation. Because stimuli of only three letters should not tax verbal working memory, subjects performed full report of the stimulus. Because we are interested in whether the identity of a letter is correctly perceived, we consider report order to be unimportant; a letter is counted as correctly identified if it appears at any position in the subject's report.

We specify location as distance from fixation in letter-widths, denoted R_n, with negative subscripts signifying the LVF. For example, for a trigram centered at fixation, the first letter falls at R_{-1}, the second at R_0, and the third at R_1. In the present experiment, a trigram can be centered at any location from R_{-4} to R_4. The trigram is presented for ~67 ms (duration titrated by subject), and immediately masked by a string of hash marks extending from R_{-6} to R_6, inclusive.

We consider the effect of position on accuracy when retinal location (and therefore acuity) is held constant. Because we will compare accuracy patterns across reading directions, we use a notation for position that is independent of whether we are considering a right-to-left (RL) language or a left-to-right (LR) language. P_L denotes the leftmost letter of a trigram (the initial letter in an LR language or the final letter in a RL language), P_M denotes the middle letter (the second letter), and P_R denotes the rightmost letter (the final letter in a LR language or the first

letter in a RL language). For example, P_R at R_{-1} means that the letter at R_{-1} is the rightmost letter of the trigram; hence two letters occurred to its left, namely P_M at R_{-2} and P_L at R_{-3}. The present experimental design provides accuracy data for all three positions at each location from R_{-3} to R_3, inclusive.

Now we consider the SERIOL predictions for a LR language. The model assumes that Letters fire briefly in sequence, and the timing of firing is determined by the amount of input from the Feature level, which is determined by the locational gradient. The location/position combination that yields the maximal Feature activation (and the earliest firing of Letters) is an LVF P_L, because its Features receive strong bottom-up excitation and no unidirectional inhibition. Let t_0 denote the time, relative to stimulus onset, at which a Letter representing an LVF P_L normally fires. If the mask begins to take effect only after t_0, an LVF P_L should always be correctly recognized because activation of the corresponding Letter is complete before the mask has any influence. As discussed below, we time the mask so that its effect presumably begins just after t_0.

We assume that the mask progressively degrades the Feature representations of the letter stimuli over time. This assumption is in line with evidence that the neural representation of a mask progressively inhibits the neural representation of the previous stimulus over a period of ~50 ms (Keysers & Perrett, 2002). Because of the mask timing, a Letter representing a letter that is not an LVF P_L will not yet have fired when the mask begins to have an effect on the Feature representations of the letters. The later such a Letter would normally fire, the less likely it will be able to fire before Feature degradation prevents its firing.

Hence accuracy should decrease with decreased Feature activation level, because the probability that the mask will inhibit the Features before the corresponding Letter can fire is increased. Recall that the activation level within the Feature area is determined by the three transformations that produce the locational gradient: (1) Stronger bottom-up (Edge-to-Feature) excitation in the RH than LH. (2) Strong unidirectional lateral inhibition within RH Features; (3) Cross-hemispheric inhibition from RH to LH Features. We consider implications of each of these assumptions.

The stronger bottom-up excitation to the RH implies that a P_L in the LVF should be perceived better than a P_L at the same distance from fixation in the RVF, because the Features of LVF/RH P_L receive strong bottom-up excitation (and no unidirectional lateral inhibition), while the Features of an RVF/LH P_L receive weaker bottom-up excitation (and no unidirectional lateral inhibition). That is, SERIOL generally predicts $(P_L, R_{-|n|}) > A(P_L, R_{|n|})$, where A denotes accuracy. In particular, with the proposed mask timing, accuracy for an LVF P_L should be at or near ceiling, while accuracy for an RVF P_L should decrease with increasing eccentricity due to the decreasing locational/acuity gradient.

The unidirectional lateral inhibition means that the Features of an LVF letter are inhibited by the Features of letters to its left. Hence, at a given location $R_{n<0}$, a P_L should not receive such inhibition, a P_M should receive inhibition from the letter to its left, and a P_R should receive inhibition from the two letters on its left. Hence the prediction for a given $R_{n<0}$ is $A(P_L) > A(P_M) > A(P_R)$.

The cross-hemispheric inhibition implies that that LVF letters can affect the perceptibility of letters that do not fall entirely within the LVF (i.e., the central and RVF letters). A central letter should show the same pattern as an LVF letter, with decreasing accuracy as the number of letters to the left increases. At R_1, a P_R should yield lower accuracy than a P_L or a P_M, because only a P_R entails the presence of an LVF letter that can drive cross-hemispheric inhibition. Position should have no effect within R_2 or R_3, as none of the trigram's letters fall within the LVF. The SERIOL predictions are summarized in Table 1.

We assumed that it would be possible to obtain the desired mask timing by titrating exposure duration per subject to yield a fairly high overall accuracy (of ~70%). This should force ceiling-level accuracy for some conditions (namely an LVF P_L), while preventing ceiling-level accuracy for all conditions. A target accuracy of 70% did successfully meet these criteria.

Location	Prediction	Reason
$R_{n<0}$ vs $R_{n>0}$	$A(P_L, R_{-\|n\|}) > A(P_L, R_{\|n\|})$	Stronger bottom-up excitation for LVF/RH
$R_{n<0}$	$A(P_L) > A(P_M) > A(P_R)$	Unidirectional inhibition
R_0	$A(P_L) > A(P_M) > A(P_R)$	Cross-hemispheric inhibition
R_1	$A(P_L) = A(P_M) > A(P_R)$	Cross-hemispheric inhibition only for P_R
R_2, R_3	$A(P_L) = A(P_M) = A(P_R)$	No unilateral or cross-hemispheric inhibition

Table 1: Summary of SERIOL predictions for the English trigram experiment.

3.1 Experiment 1

Participants

17 right-handed adult subjects served as subjects. Subjects were native English speakers, primarily undergraduate students, 18-23 (average 20.3) years old. All gave informed consent and were paid for their participation.

Stimuli

The stimuli consisted of 108 consonant trigrams, which were orthographically illegal and did not form recognizable acronyms. Consonant strings were used in order to minimize lexical and phonological processing. Trigrams were composed of all consonants of the Roman alphabet, expect Y and Q (because Y can be a vowel and Q is visually similar to the vowel O). Stimuli were presented in upper-case Courier New font in white on a black background on an LCD monitor. Each trigram subtended 1.5°.

Design

A trigram could be presented at any one of nine retinal locations, with its center letter at any location from R_{-4} to R_4, inclusive. Because each letter subtends ~0.5°, and the retinal locations include R_0 (at 0°), R_n for $n \neq 0$ corresponds to a letter centered at $0.5n$°. For example, a trigram presented at R_3 implies that the middle letter was centered at 1.5°, the left edge of the trigram fell at 0.75°, and the right edge fell at 2.25°.

The trigrams were divided into 12 groups of nine trigrams. The trigrams within a group were presented at different retinal locations. Two of the trigram groups were used for the practice blocks, in which exposure duration was titrated for each subject (described below). The remaining ten trigram groups were used in the main experiment, wherein each subject saw each trigram centered at two different retinal locations: R_n and R_{n+4} (with wrap-around if $n + 4 > 4$). Hence, twenty different trigrams were presented at each location in the main experiment. Retinal locations of the trigrams were rotated across subjects.

A practice block consisted of 18 trials, giving two trials for each retinal location. Exposure duration for the first practice block was 33 ms. Following a practice block, overall accuracy for the block was calculated. If accuracy was < 70%, the subject performed another practice block,

with exposure duration increased by one monitor refresh-cycle (17 ms). If accuracy was ≥ 70%, the subject proceeded to the main experiment, where exposure duration was set to that of the final practice block. The same two groups of trigrams were used for each practice block, with retinal location rotated across blocks. Mean exposure duration (across subjects) was 67 ms in the main experiment.

Procedure

Viewing distance was controlled with a chin rest, averaging 56 cm. Each trial commenced with a small flashing fixation cross, which appeared for 500 ms. Immediately after the offset of the fixation cross, the trigram was presented for the subject-specific exposure duration. The trigram was immediately followed by a mask in the form of a string of hash marks covering locations R_{-6} to R_6. The mask was displayed for 50 ms, and then the subject was asked to type in the letters seen, in any order, and press the Enter key. Guessing was encouraged, and input was limited to a maximum of three letters. A letter in the trigram was scored as being correctly recognized if it appeared in any position in the subject's response. The next trial automatically began 200 ms after the Enter key was pressed. The main experiment was divided into two blocks of 90 trials each, with a rest period between blocks of minimum 90 seconds and maximum 150 seconds.

Results

The results of Experiment 1 are displayed in Figure 5. The analyses are limited to those retinal locations for which data on all three positions is available, R_{-3} to R_3. The primary analysis compares performance across VFs, omitting R_0 because it does not fall within a single VF. Retinal location is broken into two factors: VF (LVF or RVF) and Distance from fixation (1, 2, or 3 letters). For example, R_{-2} corresponds to VF = LVF and Distance = 2. Hence, the primary analysis was performed via a three-way repeated-measures ANOVA: VF (LVF, RVF) x Distance (1, 2, 3) x Position (P_L, P_M,P_R). To test the prediction that a leftmost letter should be better perceived in the LVF than RVF, we also perform a VF x Distance analysis restricted to Position P_L.

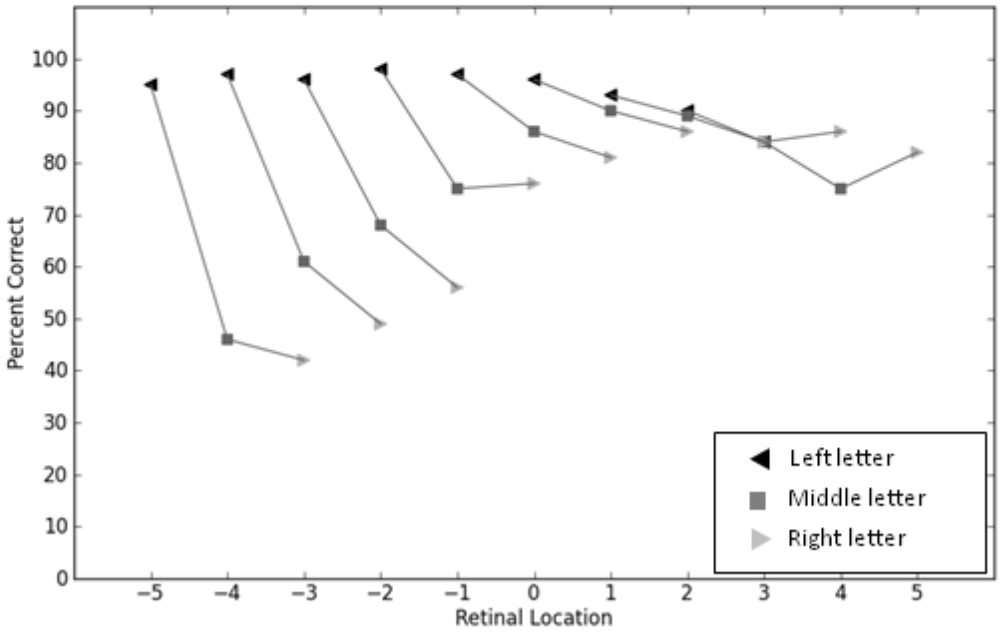

Figure 5: Data from Experiment 1, English stimuli with native English-speaking subjects. Points from the same trigram are connected.

All main effects in the primary analysis were significant: VF ($F(1,16) = 38.36$, $p < 0.001$); Position ($F(2,32) = 55.25$, $p < 0.001$); Distance ($F(2,32) = 4.03$, $p < 0.05$). The VF x Position interaction was significant ($F(2,32) = 44.62$, $p < 0.001$), as was VF x Distance ($F(2,32) = 4.03$, $p < 0.05$), but not Position x Distance ($F = 1.4$). The three-way interaction was significant ($F(4,64) = 4.47$, $p < 0.01$) and was further investigated via separate comparisons within each VF.

RVF: The main effect of Distance was significant ($F(2,32) = 3.73$, $p < 0.05$), while the main effect of Position did not reach significance ($F(2,32) = 2.52$, $p = 0.095$), but the Distance x Position interaction was significant ($F(4,l64) = 2.63$, $p < 0.05$). This interaction is due to an effect of Position at Distance 1 ($F(2,32) = 7.64$, $p < 0.01$), but not at Distance 2 ($F < 1$) or Distance 3 ($F < 1$).

LVF: The main effects of Distance ($F(2,32) = 15.72$, $p < 0.001$) and Position ($F(2,32) = 65.62$, $p < 0.001$) were significant, as was their interaction ($F(4,64) = 4.01$, $p < 0.01$). It is clear from the data that each change in Position has a large effect at each Distance; for example, the contrast of P_M vs P_R at Distance 1 is highly significant ($F(1,16) = 21.44$; $p < 0.001$). Hence the Position x Distance interaction reflects an effect of Position at all Distances, with a greater effect of Position as Distance increases.

We also performed a VF x Distance analysis restricted to P_L. The main effect of VF was significant ($F(1,16) = 12.99$, $p < 0.01$), due to higher accuracy in the LVF than RVF. The main effect of Distance was significant ($F(2,32) = 8.07$, $p < 0.01$), reflecting lower accuracy with increasing Distance. The VF x Distance interaction was significant ($F(2,32) = 5.90$, $p < 0.01$), reflecting a larger detrimental effect of Distance in the RVF than LVF.

3.2 Discussion of Experiment 1

The results are clearly in line with the predictions of the SERIOL model. For the LVF, increasing Position had a strong inhibitory effect, with $A(P_L) > A(P_M) > A(P_R)$ at all Distances. In the RVF, Position had no effect at Distances 2 or 3. Accuracy for a LVF P_L was at ceiling (~95%), independent of Distance, while accuracy for an RVF P_L decreased with increasing Distance.

We note that the data also showed the predicted patterns at R_1 and R_0. However, the experiment did not employ fixation control, due to the unavailability of an eye-tracker. Therefore, the patterns near fixation are possibly suspect, because mis-fixations of $0.25°$ to $0.75°$ (where one

letter-width = 0.5°) would have shifted a letter that was supposed to be at R_0 into the LVF or RVF, and a letter that was supposed to be at R_{-1} or R_1 to R_0. However, mis-fixations > 0.75° are rare (Van der Haegen, Drieghe, & Brysbaert, 2010), so letters at nominal Distances > 1 are highly likely to have remained in the desired VF. Van der Haegan et al. (2010) show that mis-fixations only slightly blur the true VF-specific perceptual patterns in visual word recognition. Therefore, we concentrate on the VF-specific patterns predicted by SERIOL: a strong effect of position in the LVF and little effect in the RVF, and better performance for an initial letter (a P_L) in the LVF than the RVF. Both of these predictions were strongly confirmed.

Might the full-report protocol have influenced the VF-specific patterns? Because the stimuli consisted only of three letters, transfer into and out of verbal working memory for report is assumed to be reliable. However, even if some aspect of the report process itself influenced accuracy (such as a disadvantage for the last letter), all retinal locations would be subject to this same effect. Therefore the strong VF-specific patterns cannot be an artifact of the report protocol.

4.0 Experiments with Native Hebrew Speakers

In our account of the English trigram-identification data, the strong LVF and weak RVF effects of position were taken to arise from accommodations specific to processing letter strings from left to right. Therefore, a language processed from right to left should show the opposite pattern: a strong effect of position in the RVF and little effect in the LVF. The pattern for an initial letter should also flip, with higher accuracy for a P_R (the first letter of a string in Hebrew) in the RVF than the LVF. These predictions were tested in Experiment 2, using non-word Hebrew-letter trigrams. For completeness, the same subjects were tested on the English stimuli in Experiment 3.

4.1 Experiment 2

Participants

Six right-handed adult subjects served as subjects, ages 23-44 (average 31.2 years). Subjects were bilingual Hebrew-English speakers, having Hebrew as their native language, and English as their second language. All gave informed consent and were paid for their participation.

Stimuli

108 consonant trigrams were constructed from Hebrew letters. All Hebrew letters were used and none of the trigrams formed legal roots. Stimuli were presented in white on a black background. Each trigram subtended 1.5°.

Design, Procedure

Same as Experiment 1. Average distance from screen was 56 cm, similar to Experiment 1.

Results

The data are shown in Figure 6. Initial analysis was performed via a three-way repeated-measures ANOVA: VF (LVF, RVF) x Position (P_L, P_M,P_R) x Distance (1, 2, 3).

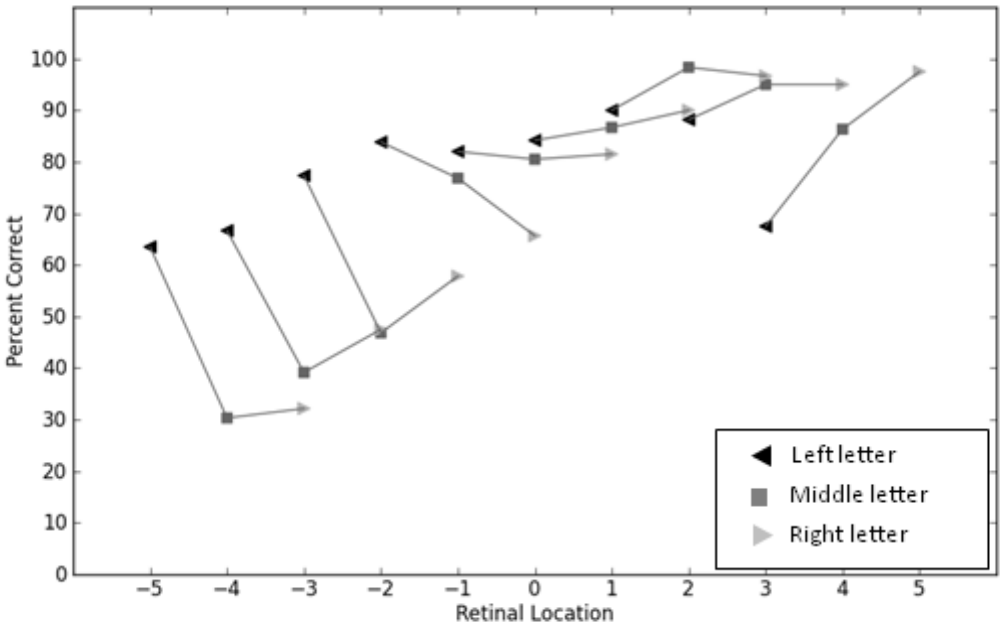

Figure 6: Data from Experiment 2, Hebrew stimuli with native Hebrew-speaking subjects.

All main effects were significant: VF ($F(1,5) = 149.74$, $p < 0.001$); Position ($F(2,10) = 12.44$, $p < 0.001$); Distance ($F(2,10) = 8.47$, $p < 0.001$). The VF x Position interaction was significant ($F(2,10) = 35.55$, $p < 0.001$), as was VF x Distance ($F(2,10) = 9.72$, $p < 0.001$), but not Position x Distance ($F = 1.1$). The three-way interaction was significant ($F(4,20) = 7.52$, $p < 0.001$) and was further investigated via separate comparisons within each VF.

RVF: The main effect of Distance did not reach significance ($F(2,10) = 2.00$, $p = 0.15$), while the effect of Position was significant ($F(2,10) = 5.71$, $p < 0.01$), as was the Distance x Position interaction ($F(4,20) = 6.05$, $p < 0.001$). This interaction is due to a strong effect of Position within Distance 3 ($F(2,10) = 9.94$, $p < 0.01$), but not Distance 1 ($F < 1$), or Distance 2 ($F(2,10) = 2.75$, $p = 0.11$).

LVF: The main effects of Distance ($F(2,10) = 13.81$, $p < 0.001$) and Position ($F(2,10) = 36.17$, $p < 0.001$) were significant, as was their interaction ($F(4,20) = 3.32$, $p < 0.05$). The interaction is due to stronger effects of Position at Distance 3 ($F(2,10) = 27.3$, $p < 0.001$) and Distance 2 $F(2,10) = 12.85$, $p < 0.001$) than at Distance 1 ($F(2,10) = 3.81$, $p = 0.06$).

We also performed a VF x Distance analysis restricted to P_R. The main effect of VF was significant ($F(1,5) = 150.00$, $p < 0.001$), reflecting higher accuracy in the RVF than the LVF. The main effect of Distance was not significant ($F < 1$), while the Distance x VF interaction was significant ($F(2,10) = 11.01$, $p < 0.001$), reflecting increasing accuracy with Distance in the RVF but decreasing accuracy with Distance in the LVF.

4.2 Experiment 3

Participants

Same as Experiment 2.

Stimuli, Design, Procedure

Same as Experiment 1.

Results

The data are shown in Figure 7. Initial analysis was performed via a three-way repeated-measures ANOVA: VF (LVF, RVF) x Position (L_1, L_2, L_3) x Distance (1, 2, 3). All main effects

were significant: VF: ($F(1,5) = 29.34$, $p < 0.001$); Position ($F(2,10) = 16.45$, $p < 0.001$); Distance ($F(2,10) = 9.67$, $p < 0.001$). The VF x Position interaction was significant ($F(2,10) = 45.90$, $p < 0.001$), but not the Position x Distance interaction ($F < 1$), or the VF x Distance interaction ($F < 1$). The three-way interaction missed significance ($F(4,20) = 2.12$, $p = 0.085$) and was further investigated via separate comparisons within each VF, in line with the other experiments.

RVF: The main effects of Distance ($F(2,10) = 3.96$, $p < 0.05$) and Position ($F(2,10) = 3.23$, $p < 0.05$) were significant, but their interaction was not ($F < 1.5$).

LVF: The main effects of Distance ($F(2,10) = 5.62$, $p < 0.01$) and Position ($F(2,10) = 63.95$, $p < 0.001$) were significant, but their interaction was not ($F < 1.5$).

We also performed a VF x Distance analysis restricted to P_R. The main effect of VF was significant ($F(1,5) = 24.14$, $p < 0.001$), due to higher accuracy in the LVF than RVF. The main effect of Distance was significant ($F(2,10) = 4.57$, $p < 0.05$), reflecting lower accuracy with increasing Distance. The VF x Distance interaction just missed significance ($F(2,10) = 3.26$, $p = 0.06$), reflecting a trend for a larger detrimental effect of Distance in the RVF than LVF.

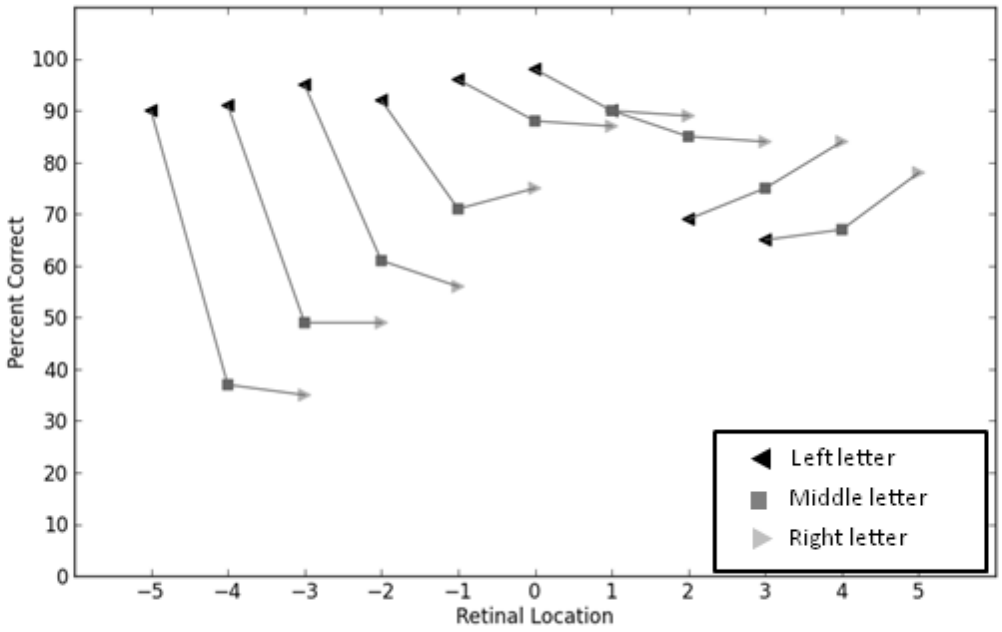

Figure 7: Data from Experiment 3, English stimuli with native Hebrew-speaking subjects.

4.3 Discussion of Experiments 2 and 3

Despite the small number of subjects (due to the difficulty of recruiting native Hebrew-speaking participants), highly significant results were achieved, reflecting the robustness of the observed patterns across subjects. To summarize, for Hebrew stimuli, Position had a strong effect at retinal locations R_{-3}, R_{-2}, and R_3, and a marginal effect at R_{-1}. For English stimuli, the Hebrew participants showed the same pattern as native English speakers, except for a disadvantage for the innermost letter in the RVF. The English results suggest that the Hebrew subjects utilized the same underlying mechanisms of orthographic processing as native English subjects, but these mechanisms were perhaps less finely tuned. Next we consider the implications of the Hebrew results.

The SERIOL prediction of strong right-to-left inhibition in the RVF/LH in Hebrew was not confirmed, as there was no effect of Position at R_1 or R_2. However, at R_3 , a P_L was less well perceived than the other positions.

The SERIOL prediction of little effect of position in the LVF/RH in Hebrew was also not confirmed, as there was a strong effect of Position at R_{-2} and R_{-3}, with P_L the best perceived, as in English. However, the LVF patterns were not exactly the same across reading direction. In English, $A(P_L) > A(P_M) > A(P_R)$ within all three LVF locations. In Hebrew, the positional effect was not reliable at R_{-1}; at the other locations, $A(P_M) = A(P_R)$. Hence, the LVF positional effect appears less systematic in Hebrew than in English.

The SERIOL prediction of higher accuracy for an initial letter (a P_R) in the RVF than the LVF was confirmed. It is of particular interest to consider the accuracy pattern for the initial letter across the RVF to fixation. Accuracy is at ceiling (~95%) for $R_{n\geq 3}$. Accuracy then drops to ~90% for R_2, ~80% for R_1, and ~65% for R_0. This pattern indicates that accuracy does not reach ceiling at R_0, even when mis-fixations are taken into account. Consider a worst-case scenario, where actual fixations are as likely on R_{-1} and R_1 as on the nominal fixation location, R_0. Using $A_O(R_n)$ to denote the observed accuracy in the present experiment for a P_R at R_n, and $A_T(R_n)$ to denote the true accuracy if fixation were controlled, we have:

$$A_O\,(R_n) = \frac{A_T(R_{n-1}) + \ A_T(\,R_n) + \ A_T(R_{n+1}\,)}{3}$$

If $A_O(R_n)$ is at ceiling, $A_T(R_{n-1})$, $A_T(\,R_n)$, and $A_T(R_{n+1}\,)$ must be at ceiling; otherwise, $A_O(R_n)$ would fall below ceiling. Therefore:

$$A_O(R_{n\geq 3}) = 95\% \ \rightarrow \ A_T(R_{n\geq 2}) = 95\%$$

Given $A_T(R_3) = 95\%$, $A_T(R_2) = 95\%$, and $A_O(R_2) = 90\%$, we can then compute:

$$A_T(R_1) = 80\%$$

Continuing these calculations for the next two locations yields:

$$A_T(R_0) = 65\% \ \text{ and } \ A_T(R_{-1}) = 50\%$$

Hence, the observed accuracies are very close to the true accuracies, and the observed accuracies are obviously at ceiling for $R_{n\geq 3}$ and below ceiling for $R_{n\leq 1}$. It is also of interest to consider an alternative scenario where $A_O(R_2)$ is instead taken to be at ceiling:

$$A_O(R_2) \approx 95\% \ \rightarrow \ A_T(R_1) \approx 95\%$$

which yields:

$$A_T(R_0) \approx 50\% \ \text{ and } \ A_T(R_{-1}) \approx 50\%$$

The important point is that $A_T(R_0)$ is well below ceiling in either case.

This result is quite surprising. Accuracy for a Hebrew initial letter is considerably lower if it is directly fixated than if it falls five letter-widths to the right! Note also that accuracy for an initial letter at fixation with *English* trigrams is at ceiling for these same subjects. Examination of the individual data reveals that every subject displayed this unexpected pattern in Hebrew; in comparing the accuracy for a P_R at R_5 versus R_0, each subject showed a disadvantage for R_0 of at least 15 percentage points. To quantify the significance of the different patterns for English and Hebrew, we performed a two-way repeated-measures ANOVA on the accuracy for an initial letter, with factors Eccentricity (0 or 5) and Language (English or Hebrew), yielding ($F(1,5) = 47.148$, $p < 0.001$) for the Eccentricity x Language interaction. Hence, the difference in the

Hebrew and English patterns is highly significant, indicating a disadvantage for an initial letter at R_0 versus R_5 in Hebrew, but no disadvantage for an initial letter at R_0 versus R_{-5} in English.

To further investigate the robustness of this surprising result, we examined the practice trials, in which exposure duration was increased from 33 ms until overall accuracy reached 70%. The shorter trial durations (than in the main experiment) should amplify the above interaction. In Hebrew, mean accuracy for a P_R was 67.5% at R_5, and 24.33% at R_0, with every subject showing a disadvantage for R_0 of at least 20 percentage points. In contrast, for the analogous comparison in the English practice trials, accuracy for a P_L was 50% at R_{-5}, and 67% at R_0; these same subjects did *not* show a disadvantage for R_0 in English. The ANOVA on the practice trials yields ($F(1,5) = 287.62$, $p < 0.00001$) for the Eccentricity x Language interaction.

In sum, we can be highly confident that the disadvantage for an initial letter at R_0 in Hebrew is a genuine effect, despite the small number of subjects. This finding places strong constraints on the nature of the underlying processing.

In summary, SERIOL predictions for Hebrew were partially confirmed. Accuracy for an initial letter was higher in the RVF than the LVF. The effect of position in the LVF was less systematic than in English, and a strong effect of position was present at one location in the RVF. However, SERIOL predicted no effect of position at any location in the LVF, and a strong effect of position at all locations in the RVF. The Hebrew data also yielded the quite unexpected result that accuracy was higher in the far RVF than at fixation.

The Hebrew data indicate that the original specifications of the lower levels of the SERIOL model are not fully correct, and should be updated. What sort of architecture would produce the observed patterns?

5.0 Revising SERIOL

We consider the implications of the trigram experiments for the architecture of the neural network supporting orthographic analysis in skilled readers. These considerations lead to the SERIOL2 model. A detailed description of how the proposed architecture would be learned during reading acquisition remains a topic for future research; we touch upon this issue briefly in the following account.

5.1 Retinotopic Letters and Abstract Letters

We first consider the Hebrew data. As discussed in Section 4.3, the accuracy pattern for the initial letter places strong constraints on the architecture of the network supporting orthographic analysis. The sharp decrease in accuracy from the far RVF to fixation indicates stronger input to letter representations encoding the far RVF than to letter representations encoding R_0. This accuracy pattern strongly suggests the existence of a weight gradient, where bottom-up weights are highest at the far RVF locations, and decrease across locations to the left.

By definition, this weight gradient occurs across a retinotopic encoding. A weight gradient would directly cause serial firing, because lower weights would delay firing onset. Therefore, we assume that serial firing originates at the level of a *retinotopic* representation in SERIOL2. (Recall that serial firing in SERIOL arose at the level of *abstract* Letter representations.) This serial firing across a retinotopic representation would be learned during reading acquisition, as a specialization for orthographic processing. Therefore, it is most natural to assume that the serial firing arises at the level of Letter, rather than Feature, representations. These factors imply the existence of Retinotopic-Letter representations, as others have proposed (Dehaene et al., 2005; Grainger & van Heuven, 2003). Retinotopic-Letters entail separate representations of a given letter at different retinal locations. For example, the stimulus "CAT" fixated on the 'A' with letters 0.5° wide would activate *T* from the set of Retinotopic-Letters centered at 0.5°, while the same stimulus fixated on the 'C' would activate a different *T* – one from the set of Retinotopic-Letters centered at 1°. We use the term RLetter as an abbreviation for Retinotopic-Letter.

In SERIOL2, a weight gradient on connections into RLetters causes the RLetters fire in sequence. All RLetters encoding a given letter connect to the same Abstract-Letter (ALetter). For example, all RLetters representing *T* (from different retinal locations) connect to the ALetter *T*. ALetters correspond to SERIOL's Letters. Activation of an RLetter causes immediate activation of the corresponding ALetter. Thus ALetters "inherit" the serial firing from the RLetters, yielding a location-invariant encoding of letter order.

The experimental data indicate that the weight gradient decreases from R_3 to R_{-5} in Hebrew (i.e., in the direction of reading), yielding the desired right-to-left firing order of the RLetters. However, the effect of position at R_3 suggests that letters to the right provide lateral inhibition, which indicates that unidirectional inhibition, rather than the weight gradient, directly supports

serial firing across locations $R_{n>3}$. This would be necessary if the weight into RLetters at R_3 attained the maximal possible value, implying that weights into $R_{n>3}$ cannot be higher than R_3. That is, weights would be non-decreasing from R_5 to R_3. The non-decreasing weights would not yield the correct firing order. Hence, unidirectional (right-to-left) lateral inhibition originating from RLetters at locations $R_{n>3}$ is necessary to induce the correct firing order, by delaying the firing of RLetters to the left. See the top diagram of Figure 8. We will continue to use the term *weight gradient* to refer to the pattern across retinal locations of excitatory weights into RLetters, even though these weights are not monotonically decreasing across all locations.

One aspect of the Hebrew data remains to be explained. For $R_{n\leq -2}$, recall that accuracy is highest for a P_L, while $A(P_M) = \mathrm{A}(P_R)$. As discussed in more detail in Section 5.2, we propose a general attentional advantage for the outermost letter (i.e., an advantage not related to orthographic processing). This general effect interacts with the lateralization of orthographic processing to the left hemisphere, yielding a much stronger advantage for an outermost letter in the LVF than in the RVF. This interaction creates a general advantage for the LVF outermost letter, which is unrelated to reading direction.

Next we consider the English data. For $R_{n\leq 0}$, accuracy for the initial letter (P_L) is at ceiling, and $A(P_L) > A(P_M) > A(P_R)$. This pattern indicates that the weight gradient takes near maximal values for the LVF/central locations. That is, weights are high and non-decreasing for $R_{n\leq 0}$, and serial firing across these locations is induced by unidirectional inhibition. For $R_{n>0}$, the experimental lack of positional effect and the decreasing accuracy with increasing eccentricity indicate that weights decrease with increasing *n*; the weight gradient produces serial firing across these locations. See the bottom diagram of Figure 8.

In English, the proposed general advantage for the LVF outermost letter aligns with the unidirectional inhibition. That is, the first letter (P_L) is at an advantage both because does not receive unidirectional inhibition, and it is an outermost letter. A second letter (P_M) is at a disadvantage both because it receives inhibition from P_L, and it is not an outermost letter. A third letter (P_R) is at an even greater disadvantage because it receives inhibition from both P_L and P_M (and it is not an outermost letter). In contrast, for Hebrew trigrams, P_M and P_R in the LVF do not

receive unidirectional inhibition; they are each at a similar disadvantage from not being an outermost letter. This explains why the LVF pattern is more graded in English than Hebrew.

The proposed patterns of weights and unidirectional inhibition are not mirror-images of each other across reading direction. Why might this be the case? Our explanation is related to the proposed source of the advantage for the LVF outermost letter, which we discuss next. After addressing this advantage, we return to the issue of weight-gradient patterns across reading directions.

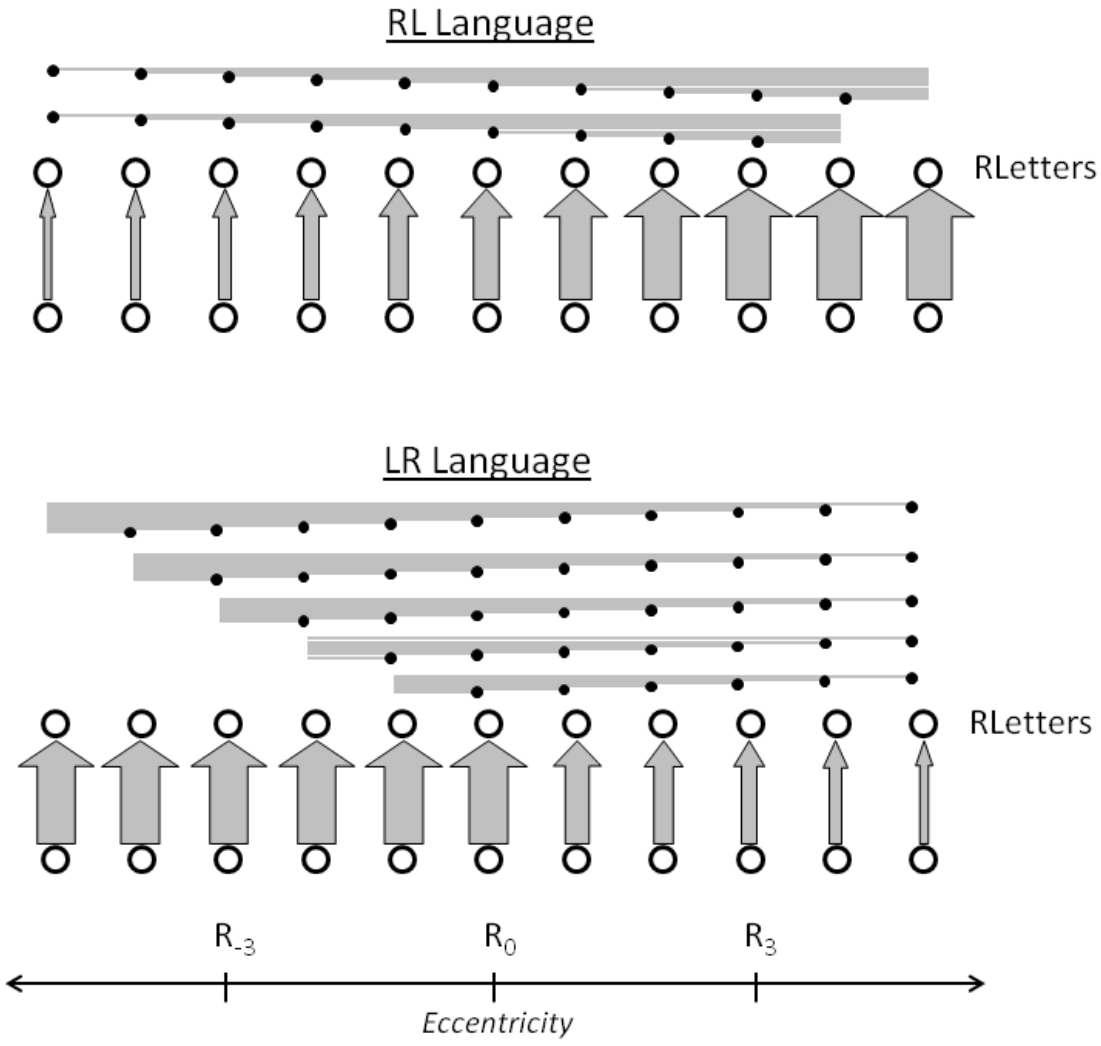

Figure 8: Proposed connectivity into and among RLetters, by reading direction. Wider arrows represent higher weights. For clarity, unidirectional inhibition is shown above the RLetters. Unidirectional inhibition originates from RLetters where the weight gradient is non-decreasing in the direction of reading. For example, considering only the excitatory weights into RLetters in Hebrew, RLetters at R_5, R_4 and R_3 would all fire at the same time because they have equivalent weights. Therefore, RLetters at R_5 and R_4 each inhibit all locations to the left, to delay their firing.

5.2 Lateralization of RLetters

We suggest that a general advantage for an outermost letter is not directly related to orthographic processing, but rather is an artifact of the experimental protocol, where a string could appear abruptly within a single VF. Studies of visual crowding have demonstrated that a distractor object on the outer side of a lateralized target is more inhibitory than a distractor object on the inner side (Petrov, Popple, & McKee, 2007). That is, the target is better perceived when it is the outermost object in the visual field. This advantage for the outermost object stems from the allocation of visual attention (Petrov & Meleshkevich, 2011). Hence we suggest that a non-specific advantage for the outermost object interacts with the letter-specific encoding mechanisms to yield the patterns observed in the trigram experiment. This attentional effect would not be a factor when the subject knows the location of the upcoming letter string, such as in normal reading.

Why then is the outer-letter advantage much stronger in the LVF than the RVF? We suggest that the answer is related to the known left-lateralization of orthographic processing. In particular, we propose RLetters representing both visual fields reside in the LH. Because the LH is primarily devoted to representing the RVF, RLetters tuned to LVF locations are situated near the representation of the foveal center (R_0) in the LH. Therefore, the LVF and central RLetters are cortically close to each other, while the central and RVF RLetters are more distant from each other. Because cells that are near each other tend to non-specifically inhibit one another (Douglas & Martin, 2004; Mariño et al., 2005), LVF RLetters inhibit each other more than RVF RLetters inhibit each other.

This general inhibition for LVF RLetters amplifies the attentional effect, as follows. Initially, the outer RLetter has a higher activity level than the inner RLetters, due to the effect of attention. Over time, the outer RLetter inhibits an inner RLetter more than an inner RLetter inhibits the outer RLetter. This difference in inhibition allows the outer RLetter to become even more highly activated than the inner RLetters, magnifying the attentional effect. Therefore, an outermost letter has more of an advantage in the LVF than the RVF.

However, this account is inconsistent with the notion that RLetters fire rapidly in sequence. If RLetters fire serially, they would not have the opportunity to inhibit each other in the proposed manner. Also, it would make sense for LVF letters to be recognized within the RH, where the

necessary feature information is directly accessible. That is, we desire LVF RLetters to be situated in the RH for letter recognition, but to be lateralized to the LH to explain the strong advantage for the outermost letter in the LVF. We require RLetters to fire serially as suggested by the Hebrew trigram data, but we desire RLetters to interact in parallel to explain the strong advantage for the outermost letter in the LVF. These conflicts can be resolved by assuming multiple layers of RLetters.

5.3 Layers of RLetters

It is well known that the cortex is comprised of layers, with bottom-up input (from a lower-level area) arriving in layer 4 (L4). In general, L4 projects to L2/3, which sends feed-forward input to higher-level areas, provides lateral connections within an area, and receives feedback from higher-level areas (Douglas & Martin, 2004). Studies of the propagation of signals between L4 and L2/3 suggest that L2/3 strongly gates the timing and degree of transmission of information from L4 (Lübke, Roth, Feldmeyer, & Sakmann, 2003).

Therefore, we assume multiple layers of RLetter representations, as follows. L4 RLetters, which are present in both the LH and RH, receive input from the Feature area. L4 RLetters recognize letters. L4 RLetters from both hemispheres project to L3 RLetters lateralized to the LH. The L3 RLetters act as a buffer to support serial firing of L2 RLetters. L2 RLetters provide the output of the RLetter area, which is the input to the ALetter area. We assume one-to-one excitatory connections between RLetter layers. For example, an L4 RLetter tuned to a given letter and location strongly excites one L3 RLetter, which necessarily then encodes the same letter and location.

L4 RLetters fire in parallel and L2 RLetters fire strictly in sequence. L3 RLetters mediate the transition from a parallel to a serial encoding. That is, the firing of different L3 RLetters can overlap in time (parallel encoding), while the latency (time of first spike) of different L3 RLetters can vary (serial encoding). The weight gradient occurs on L4 → L3 connections, non-specific lateral inhibition operates among L3 RLetters tuned to LVF/central locations, and unidirectional inhibition operates from L3 to L2 RLetters. We also assume strong feedback inhibition from L2 to L3 RLetters (within a location), to assure strictly serial firing across L2 RLetters. This architecture satisfies our requirements. LVF letters are recognized in the RH, and

RVF letters are recognized the LH. L3 RLetters tuned to LVF/central locations are cortically near each other in the LH, and can interact during the process of serial activation of L2 RLetters.

Letter recognition can be considered to be both parallel and serial, as L4 RLetters become active simultaneously, but L2 RLetters sequentially pass letter information to higher-level areas. This duality is consistent with conflicting evidence from visual-word recognition studies of positional effects as a function of exposure duration. Accuracy is at chance at all letter positions for an exposure of 18 ms, but above chance for all positions for an exposure of 24 ms, suggesting parallel processing (Adelman, Marquis, & Sabatos-DeVito, 2010). However, accuracy decreases across letter position (Adelman et al., 2010; Adelman, 2011), suggesting serial processing. The step from chance to above-chance performance would reflect parallel activation of L4 RLetters. However, L4 RLetters must activate the corresponding L2 RLetters for letter-identity information to be accessible to higher-level areas. Because L2 RLetters spike serially, the interval between L4 and L2 activation increases with increasing string position. A longer interval increases the probability that the mask will inhibit the L4 RLetter before it can activate the correct L2 RLetter. Therefore, accuracy decreases across position.

All cortical areas of the brain consist of multiple layers. In SERIOL2, we only explicitly model multiple layers in the RLetter area; other SERIOL2 areas specify the output-layer representations (L2). We focus on the multiple layers of the RLetter area because we propose that the important parallel-to-serial transformation occurs between these layers.

Where in the brain would RLetters reside? Szwed et al. (2011) identified letter-specific fMRI activity (stronger activation for letters than scrambled letters) in bilateral V1/V2 and bilateral V3/V4. Letter-specific activity significantly interacted with hemisphere in V3/V4 (stronger effect in the LH than the RH), but not in V1/V2. Hence we place letter Features in V1/V2 and RLetters in V3/V4. The bilateral L4 RLetters explain the observed bilateral letter-specific effect in V3/V4, while the lateralization of L2/3 RLetters to the LH explains the stronger letter-specific effect in left than right V3/V4.

How far apart in time would successive L2 RLetters fire? To fit the trigram data, the SERIOL2 simulations presented in Section 7 yielded intervals of 5 to 10 ms between successive L2 RLetters. This timing is somewhat faster than proposed for SERIOL (~16 ms between successive

Letters). Is the SERIOL2 timing realistic? A recent study of the encoding of item order in visual working memory is relevant to this issue (Siegel, Warden, & Miller, 2009). An analysis of correct trials showed that neurons representing the identity of the first item spiked 57° earlier, relative to a 32 Hz oscillation in Local Field Potential, than neurons representing the identity of the second item. Translating this phase difference to milliseconds yields 57°/360° * 1000 ms/32 ≈ 5 ms. Inclusion of incorrect trials into the analysis abolished the phase difference, indicating that behavioral performance was related to spike timing. Hence, this study demonstrates that item order is encoded by spike timing on the scale of 5 ms / item, consistent our proposal that letter order is represented by spike timing on the scale of 5-10 ms / letter.

5.4 Weight Gradients

The proposal for cortical RLetter layout also contributes to the explanation of why the weight gradients for RL and LR languages are not mirror images of each other. In the following, we discuss the pattern of excitatory weights into L3 RLetters (which are lateralized to the LH). In particular, we assume that a connection is comprised of multiple synapses, and that the weight on a connection is the product of the number and strength of the synapses. We first consider the patterns of synaptic strengths and number of synapses, and then the resulting weight gradients. For brevity, we refer to L3 RLetters simply as RLetters for the remainder of this Section.

We propose that RLetters that are cortically near each other develop similar excitatory synaptic strengths, presumably due to perceptual learning under the spatial attentional patterns involved in reading acquisition. (The details are a subject of future research.) Therefore, synaptic strengths are equivalent across $R_{n\leq 0}$. We also assume that synaptic strength is driven to the maximal value for RLetters tuned to the location at which the initial letter of a string usually falls during normal text reading (i.e., $\sim R_{-3}$ for an LR language, and $\sim R_3$ for an RL language), and that synaptic strength decreases as the *cortical* distance from RLetters tuned to this location increases. Hence, for an LR language, synaptic strength is maximal at R_{-3}, and therefore maximal for all $R_{n\leq 0}$; synaptic strength decreases with increasing n for $R_{n\geq 0}$. For an RL language, synaptic strength is maximal at R_3, and decreases away from R_3; synaptic strength is uniformly low at $R_{n\leq 0}$.

Next we consider number of synapses. RLetters tuned to the LVF receive input from the RH, while RLetters tuned to the RVF received input from the LH. For LVF RLetters, the number of synapses is taken to decrease with increasing eccentricity of the tuned location, due to a

decreasing density of cross-hemispheric fibers with increasing distance from the vertical meridian (Van Essen, Newsome, & Bixby, 1982). For RVF RLetters, the number of synapses is taken to be unaffected by eccentricity due to strong connectivity within the LH.

These assumptions yield the patterns portrayed in Figure 9. Next we examine the resulting weight gradients for each reading direction.

For an RL language, we consider the pattern from R_5 toward R_{-5} (i.e., in the direction of reading). Synaptic strength increases from R_5 to R_3 and decreases from R_3 to R_0, while number of synapses is constant from R_5 to R_0. Synaptic strength is constant from R_0 to R_{-5}, while the number of synapses decreases. As a result, connection weights increase from R_5 to R_3, and then decrease from R_3 to R_{-5}. Hence, the weight gradient yields the correct order of firing from R_3 to R_{-5} .

For an LR language, we consider the pattern from R_{-5} toward R_5. Synaptic strength is constant from R_{-5} to R_0, while number of synapses increases. Synaptic strength decreases from R_0 to R_5, while number of synapses is constant. As a result, connection weights increase from R_{-5} to R_0, and decrease from R_0 to R_5. Hence, the weight gradient yields the correct order of firing from R_0 to R_5.

We propose that *inhibitory* weights are not subject to the same constraint of equivalent synaptic strengths at nearby cortical locations, and that learned unidirectional inhibition originates from locations where the weight gradient is non-decreasing. (Again, the details are a topic of future research.) For an RL language, unidirectional inhibition originates from R_5 and R_4. For an LR language, unidirectional inhibition originates from R_{-5} to R_{-1}.

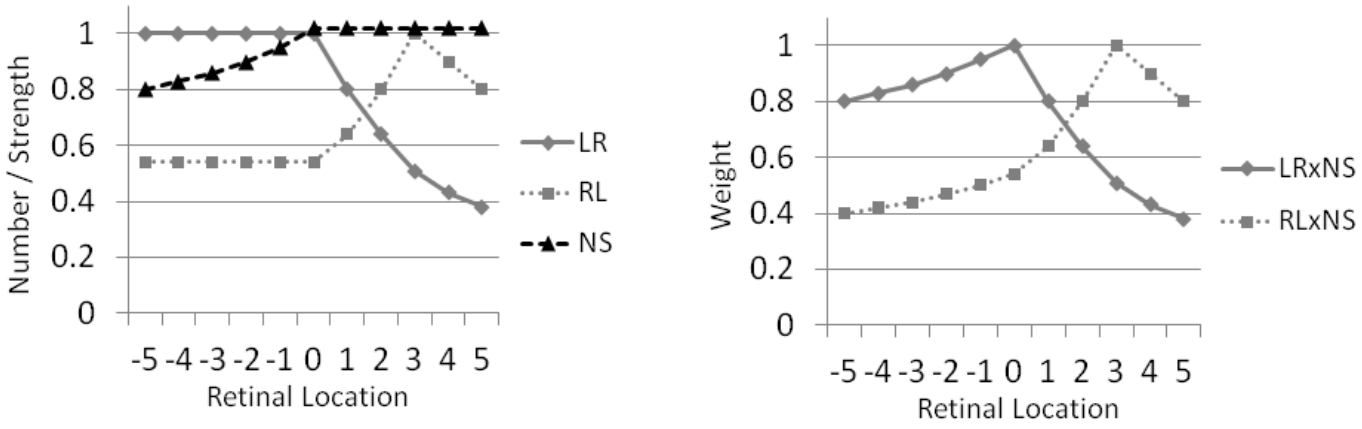

Figure 9: Proposed composition of weight gradients. The graph on the left illustrates the number of synapses (NS), as well as the synaptic strengths for LR versus RL languages. A value of 1 represents the maximal number of synapses, or the maximal synaptic strength. As shown in the graph on the right, the weight at each retinal location is the product of the number of synapses and the synaptic strength, yielding the weight gradients for each reading direction. Note that the weights illustrated in Figure 8 should be taken as a first approximation of these weight gradients. In particular, the equivalent maximal weights of Figure 8 (for $R_{n\geq 3}$ in RL, and for $R_{n\leq 0}$ in LR) are replaced here by weights that increase from 0.8 to 1 in the direction of reading.

5.5 Summary of the Lower Levels of SERIOL2

We conclude the discussion of letter representations by summarizing the architecture of the lower levels of SERIOL2, and comparing to SERIOL. In SERIOL2, a bilateral Feature area connects to bilateral L4 RLetters, and all L4 RLetters connect to L3 RLetters lateralized to the LH, which connect to L2 RLetters (lateralized to the LH). L2 RLetters connect to ALetters. SERIOL2 re-specifies the way in which the serial encoding of letter order is induced. In SERIOL, differential bottom-up weights and unidirectional inhibition at the (retinotopic) Feature level create an activation gradient (i.e. the locational gradient), which induces serial firing at the (abstract) Letter level. In SERIOL2, differential bottom-up weights and unidirectional inhibition at the RLetter level directly cause serial firing across RLetters. The ALetters then "inherit" the serial encoding from the RLetters. Hence, SERIOL2's Feature, RLetter, and ALetter areas replace SERIOL's Feature and Letter areas. See Figure 10.

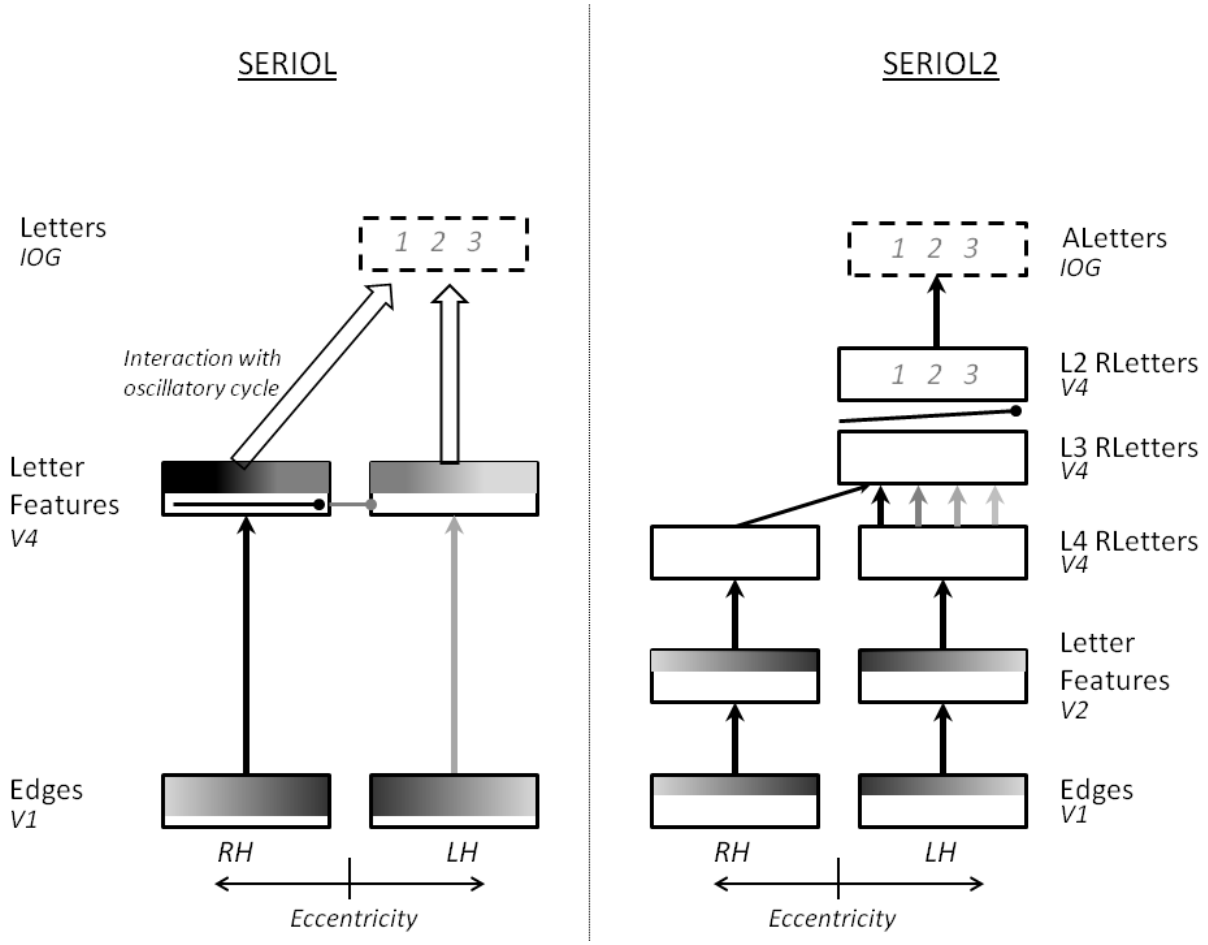

Figure 10: Comparison of the SERIOL and SERIOL2 models for a LR language, using the same notation as Figure 4. The lower areas of SERIOL are repeated from Figure 4. The thinner grey-level gradients for SERIOL2 indicate that activation gradients are not integral to this model; SERIOL2 does not employ the locational gradient, nor the oscillatory mechanism. In SERIOL, differential weights occur on Edge→Feature connections (i.e., stronger weights in the RH), and unidirectional lateral inhibition operates within the Feature area; these mechanisms yield the locational gradient, which induces serial firing in the (abstract) Letter area. In SERIOL2, differential weights occur on the L4→L3 RLetter connections (i.e., the weight gradient), and unidirectional lateral inhibition operates between L3 and L2 RLetters; these mechanisms directly induce serial firing across L2 RLetters. The ALetter area then inherits the serial encoding.

5.6 Open-Bigrams

SERIOL2 remains essentially the same as SERIOL above the ALetter area. That is, ALetters connect to Open-Bigrams on the ventral pathway and to Graphemes/Phonemes on the dorsal pathway. Open-Bigrams in SERIOL2 are activated in the same manner as SERIOL – via the order of firing of the constituent ALetters. However, SERIOL2 specifies several additional assumptions about Open-Bigrams, to increase explanatory capacity.

The first new assumption is that an Open-Bigram continues to fire once it is activated; after the last Open-Bigram is activated, all Open-Bigrams fire in parallel. The second new assumption is the existence of feedback connections from VWFs to Open-Bigrams. These two assumptions allow VWFs to affect Open-Bigram activity, which in turn affects VWF activity. This feedback explains the known facilitative effect of a dense orthographic neighborhood (Whitney, 2011), as addressed in more detail in Section 8.1.

The third new assumption is the generalization of Edge Bigrams. SERIOL included Edge Bigrams that were only activated by exterior letters. For example, "art" would activate **A*, but "rat" would not. We now assume graded activation of Edge Bigrams, like any other Open-Bigram; "rat" would induce a medium activation level of **A* (as well as *A**). This is consistent with recent evidence that letter position is encoded relative to word edges (Fischer-Baum, Charny, & McCloskey, 2011).

The fourth new assumption is related to the implementation of graded activation. Recall that the activation level of an Open-Bigram is taken to decrease as the interval between the spiking of the constituent ALetters increases. With spiking neurons, two different mechanisms have been proposed for how graded activations could be realized. One mechanism is based on graded firing rates of individual neurons. The other mechanism employs a pool of neurons with similar tuning, where each neuron is either active or not; activation level corresponds to the number of active neurons in the pool. Assuming that active neurons fire near synchronously, this latter mechanism has the advantage that information about activation level is available at the time scale of a few milliseconds, whereas extraction of this information from a rate coding requires integration over a longer time period.

SERIOL2 posits the latter mechanism for graded activation of Open-Bigrams. That is, a pool of Open-Bigrams exists for each open-bigram. The members of the pool have different tolerances for the temporal proximity of the constituent ALetters. For example, multiple neurons would detect *X*-then-*Y* (an *XY* Open-Bigram), but the maximal allowable time between the firing and *X* and *Y* would vary among these neurons. Suppose that one *XY* neuron fires if *Y* spikes at most 7 ms after *X*, while another *XY* neuron fires if *Y* spikes at most 15 ms after *X*. If *X* and *Y* spike 5 ms apart, both *XY* neurons fire; if *X* and *Y* spike 10 ms apart, only the latter *XY* neuron can fire. Thus total activity in the *XY* pool decreases as the interval between the firing of *X* and *Y* increases, yielding graded activation of *XY*. This implementation of graded Open-Bigram activity contributes to the explanation VF-specific orthographic-neighborhood effects, as discussed in Section 8.1.

5.7 VWFs

In SERIOL2, we formalize several assumptions about VWF activation. We assume inhibitory connections between VWFs, and extended settling dynamics. In particular, we assume that Open-Bigrams have activated multiple VWFs by ~200 ms post-stimulus. The VWFs compete with each other until the network settles, and a winning VWF emerges by 600 ms.

This is consistent with an Event-Related Potential study showing that effects of orthographic-neighborhood density (OD) are strongest from ~250 to ~400 post-stimulus (reflecting early activation of multiple VWFs) and have disappeared by 600 ms, while lexical-frequency effects are strongest from ~400 to ~600 ms post-stimulus (reflecting lexical selection of the winning VWF) (Vergara-Martínez & Swaab, 2012). During the period from ~400 to ~500 ms, the OD effect is stronger and more posterior for low-frequency than high-frequency words. This interaction of the OD effect with frequency indicates that lexical selection is largely complete for high-frequency words (i.e., lexical competitors have already been inhibited, yielding little effect of OD), while lexical selection is ongoing for low-frequency words. The posterior distribution suggests an effect on prelexical orthographic representations (Vergara-Martínez & Swaab, 2012), consistent with our assumption of recurrent excitation between VWFs and Open-Bigrams. Note that these timings are for experimental conditions, where the subject has no information about the identity of an upcoming stimulus. During normal reading, the process of lexical selection would be speeded by syntax, semantics, and parafoveal preview.

Increased word length is not expected to cause longer reaction times, because the lexical network does not settle immediately after activation of the final Open-Bigram. In fact, increased word length yields greater total Open-Bigram activity, which would cause faster VWF activation and *decreased* settling times (even under serial activation of Open-Bigrams), unless connection weights are normalized (Whitney, 2011). Hence we assume that Open-Bigram→VWF weights are weaker for longer words. Next we concisely specify SERIOL2.

6.0 Specification of the SERIOL2 Model

6.1 Features

The Feature area corresponds to bilateral V1/V2. Features are active in parallel and connect to Retinotopic-Letters.

6.2 Retinotopic Letters

The Retinotopic-Letter area corresponds to bilateral V3/V4. Layer 4 Retinotopic Letters (L4 RLetters) recognize letters in parallel. All L4 RLetters connect to L3 RLetters that are lateralized to the LH. L3 RLetters connect to L2 RLetters (also lateralized to the LH). Serial firing occurs across L2 RLetters. The time between successive L2 RLetters is taken to be 5 – 10 ms.

L2/3 RLetters tuned to the LVF develop at the representation of the foveal center (in the LH). Therefore LH RLetters tuned to center and LVF are all cortically near each other. As a result, LVF/central LH RLetters non-specifically inhibit one another.

Connection weights between L4 and L3 RLetters are the product of two factors. One factor is the synaptic strength learned due to top-down attentional modulation, where the central and LVF L3 RLetters all develop similar synaptic strengths due to their cortical proximity. The other factor is a decreasing number of synapses from RH L4 RLetters with increasing eccentricity. The resulting weight gradients are shown in Figure 9.

Where the weight gradient cannot directly create serial firing, unidirectional inhibition among the L2/3 RLetters is learned. For a LR language, an LVF L3 RLetter inhibits all L2 RLetters tuned to locations to the right. For a RL language, L3 RLetters tuned to locations $> 1°$ inhibit all L2 RLetters tuned to locations to the left.

6.3 Abstract Letters

The Abstract-Letter area corresponds to the left IOG. All L2 RLetters encoding a given letter have strong connections to the corresponding Abstract-Letter (ALetter), such that the ALetter immediately spikes if any one of these L2 RLetters spikes. After a single spike or burst, an ALetter is quiescent unless it is re-activated by a different L2 RLetter. ALetters connect to Open-Bigrams on the ventral pathway and Graphemes on the dorsal pathway. We focus on the ventral pathway.

6.4 Open-Bigrams

The Open-Bigram area corresponds to left pOTS. We define an *XY* Pool as set of *n* Open-Bigrams XY_i, where XY_i fires if ALetter *Y* spikes within t_i ms of ALetter *X*. We take t_i to be in the range [~5 ms, ~25 ms]. The total activity in the Pool decreases as the interval between the firing of *X* and *Y* increases. The same dynamics apply to Edge Open-Bigrams, where *X* corresponds encodes a space before the first letter of a string, or *Y* a space after the last letter.

An Open-Bigram continues to spike once activated. An Open-Bigram also receives top-down excitation from VWFs. Convergent top-down excitation (from many VWFs) can cause a quiescent Open-Bigram to begin spiking.

6.5 Visual Word Forms

The VWF area corresponds to left mFUS. An Open-Bigram has a high connection weight to a VWF if that open-bigram is present in the word. Connection weights are weaker for longer words. Open-bigrams activate multiple VWFs. VWFs inhibit each other, leading to an extended settling process at the VWF level.

7.0 Simulations

We now present simulations showing that SERIOL2 can yield serial firing of L2 RLetters and can explain the observed trigram patterns. Because the proposed processing depends on spike timing with millisecond precision, the simulations utilize a spiking neural network of leaky integrate-and-fire neurons. A simulation requires simplifying assumptions and choice of particular parameters. Our goal was to construct the minimal network that would illustrate our proposals.

The network consists of four layers: Features, L4 RLetters, L3 RLetters, and L2 RLetters. Each layer has excitatory one-to-one connections into the next layer. The simulation includes 11 locations representing R_{-5} to R_5, corresponding to those utilized in the trigram experiments. For simplicity, each location in each layer is comprised of a single simulated neuron (called a *node*), which is meant to represent a group of Features, or the neural assembly corresponding to the correct RLetter (in layers L2, L3, and L4). Nodes in the Feature layer are forced to spike at a high rate as a Poisson process, providing the input to the network.

Connection weights were set to embody the SERIOL2 proposals for a skilled reader. Excitatory Feature→L4 weights are constant across location. Excitatory L4→L3 weights implement the weight gradients illustrated in Figure 9. Excitatory L3→L2 weights are also constant. Inhibitory L3→L2 connections (across different locations) implement the unidirectional inhibition illustrated in Figure 8. All L3 nodes at $R_{n\leq 0}$ weakly inhibit each other, to implement the proposed non-specific lateral inhibition due to cortical proximity.

Multiple L3 spikes are required for an L2 node to spike. An L3 node possesses an excitatory self-connection to speed re-spiking. The rate of L3 spiking is influenced by the L4→L3 excitatory weight. L2 nodes send one-to-one inhibitory connections back to L3 nodes. When an L2 node spikes, it strongly inhibits its L3 node, preventing further L2 excitation or inhibition by that L3 node. See Figure 11 for a diagram of the network structure.

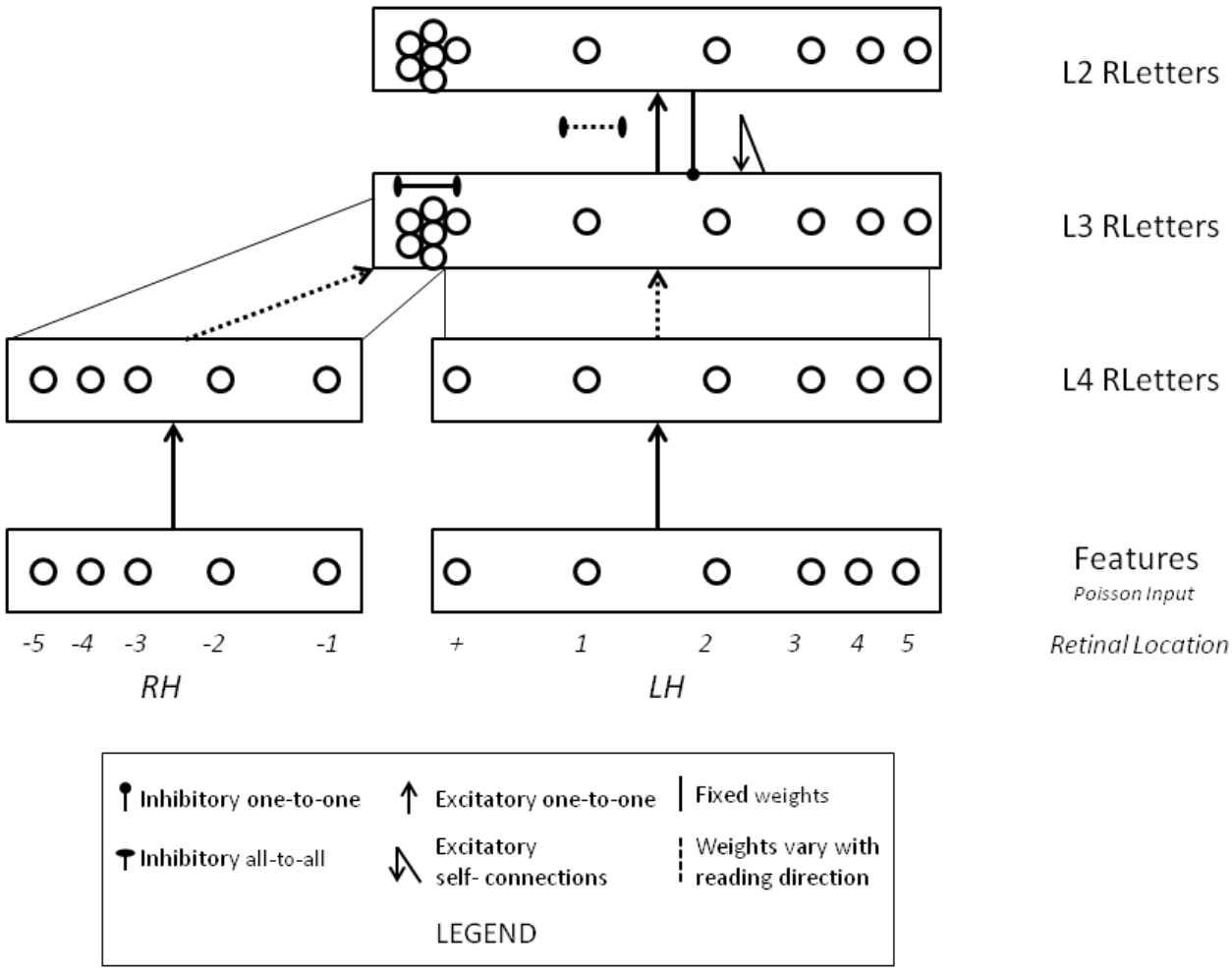

Figure 11: Architecture of simulated network. Each circle represents a spiking neuron. Spacing between neurons is shown to suggest the cortical layout underlying the pattern of non-specific lateral inhibition among L3 RLetters. The L4→L3 excitatory connections implement the weight gradients. L3→L2 inhibitory connections are shown as all-to-all to indicate that any pattern of connection is potentially possible; in particular, these connections implement the unidirectional inhibition.

Parameters were hand tuned to yield the desired serial firing pattern across the L2 nodes in simulations of normal string processing, and to replicate the accuracy patterns of Experiments 1 and 2 in trigram simulations. The parameter values are given in the Appendix.

The goal of a simulation of normal string processing is for the L2 nodes to serially fire from left to right for a LR language, and from right to left for a RL language. In a RL versus LR simulation, all parameters are the same except for the L4→L3 excitatory connections (weight gradient), and the L3→L2 inhibitory connections (unidirectional lateral inhibition). The presence of an input letter at a given location is simulated by turning on the Feature node for that location. For simulations of normal string processing, we simulate six-letter strings presented at a range of locations. For the LR simulations, the location of the initial letter (start location) ranged from R_{-5} to R_0; for the RL simulations, start location ranged from R_5 to R_0. The simulation of different start locations demonstrates that serial firing is maintained for strings spanning different locations, and illustrates the varying patterns across reading direction.

The upper graph of Figure 12 displays the average spiking time at each location for the LR simulations. The firing time at a given location is shifted upward as start location moves to the left, due to the increasing unidirectional inhibition from L3 →L2 RLetters.

The lower graph of Figure 12 displays the results for the RL simulations. Firing time at a given location is fairly constant if the initial letter falls at $R_{n\leq 3}$, because the fixed weight gradient determines firing time. Because the weight gradient peaks at R_3, L3 RLetters at $R_{n>3}$ inhibit all L2 RLetters to the left. Hence, for start locations $R_{n>3}$, firing times are shifted upwards.

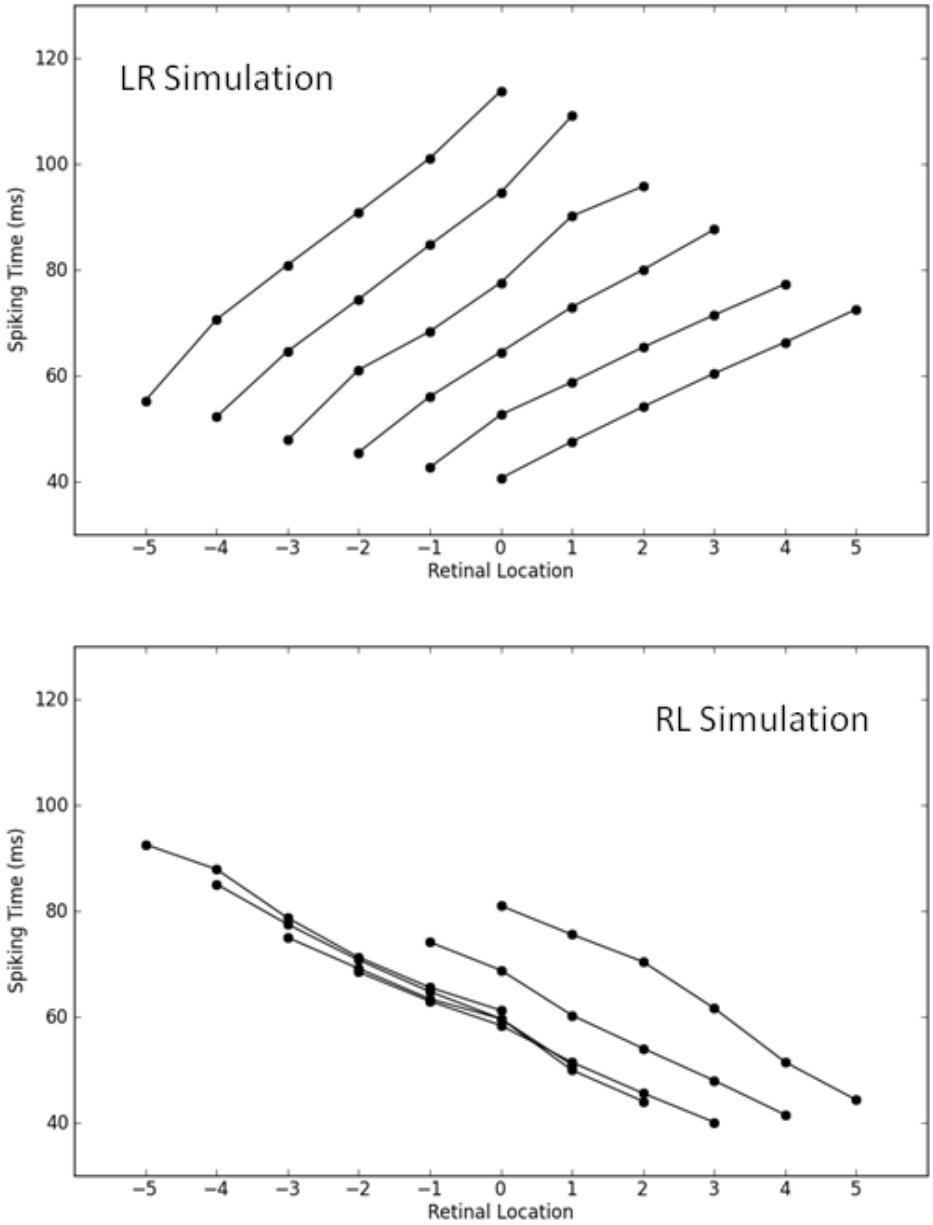

Figure 12: Spiking times of L2 RLetters for simulations of normal string processing, averaged over 10 runs for each start location. Points from the same start location are connected.

We now consider simulations of the trigram experiments. It is unlikely that a simple simulation will yield results that exactly match the data of Experiments 1 and 2. Therefore, the objective of the trigram simulations is to replicate the major characteristics of the experimental patterns. In particular, we wish to achieve the following goals: 1) For both LR and RL simulations, a stronger effect of position in the LVF than the RVF. 2) For LVF locations, $A(P_R) < A(P_M)$ for the LR simulation, but not the RL simulation. 3) For the LR simulation, ceiling-level accuracy in the LVF and center, and reduced accuracy in the far RVF. 4) For the RL simulation, ceiling-level accuracy in the far RVF, and reduced accuracy at the center and LVF.

The trigram simulations were performed as follows. 200 runs were performed for each trigram location. For each run, a "mask time" was selected, ranging from 30 to 70 ms. A simulation proceeded normally until the mask time was reached. Then, the Feature firing rate was linearly decreased to 0 Hz over an interval of 40 ms. Inhibition was also injected into L2 nodes (via a Poisson process) starting at the mask time, to simulate direct inhibitory influences of the mask. The variation in mask times is meant to simulate varying degrees of readiness of the visual system to begin processing the stimulus, related to attention level and phase of ongoing oscillatory activity (Besserve et al., 2008). Hence an early mask time (e.g., 30 ms) corresponds to low readiness, which delays the onset of letter processing, yielding less processing time prior to the mask.

The known effect of greater attention for an outermost item was simulated by setting the Feature firing rate for the two inner locations of a trigram to lower rates than normal. If the trigram was centered at fixation, this reduction was applied to the outer two letters to simulate increased attention to the fixation point.

Otherwise, the normal and trigram simulations used the same parameters. Accuracy for a given position/location is the number of runs in which the corresponding L2 node spikes, divided by the total number of runs (= 200). Variability in L2 spiking across runs (for the same trigram location) arises from the differing mask times, and the randomness of the Poisson processes governing the Feature and mask-inhibition firing rates. The results of the trigram simulations are presented in Figures 13 and 14.

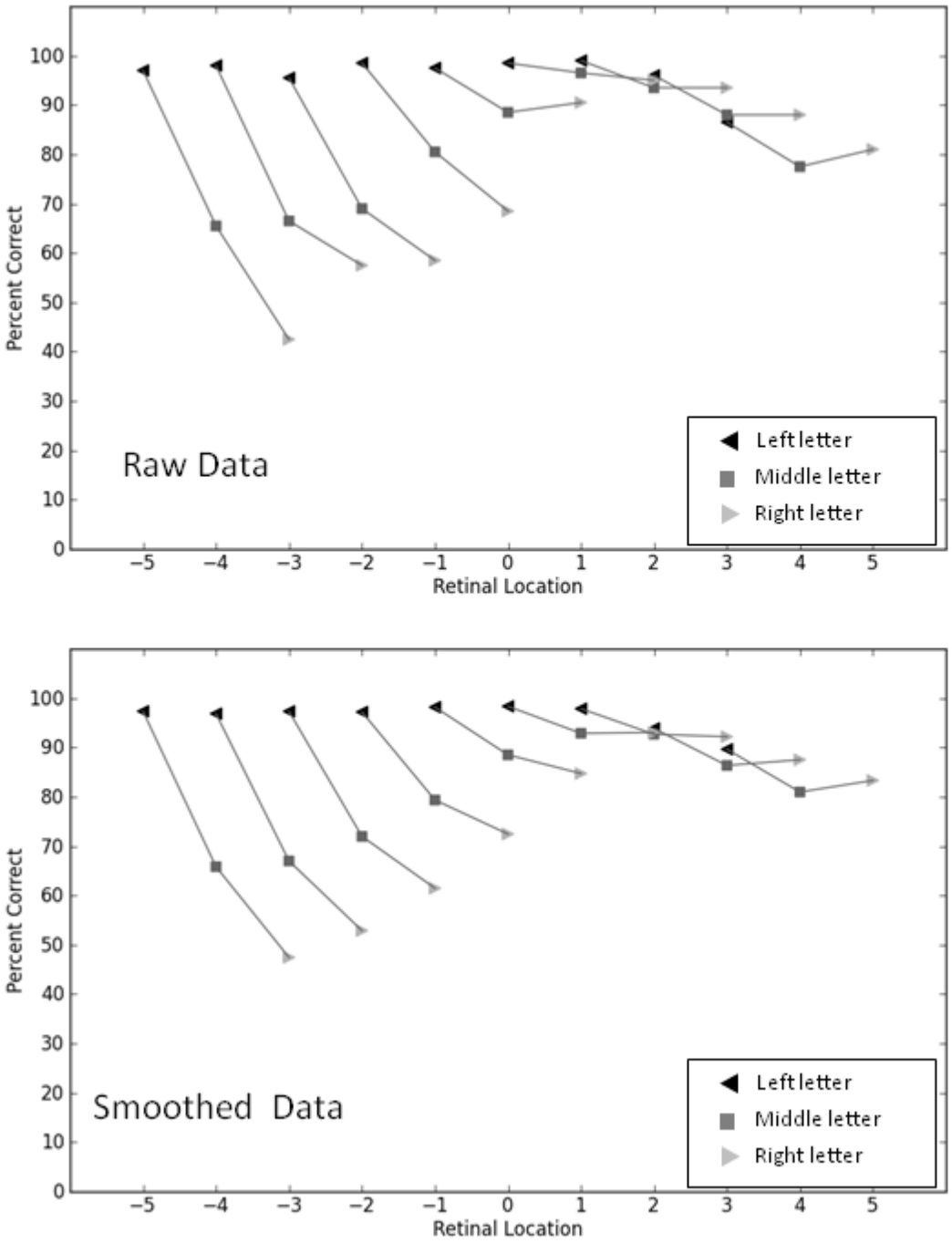

Figure 13: Trigram simulations for a LR language. The top panel displays the raw results. To illustrate the possible effect of mis-fixations, the bottom panel displays the accuracy at each location as the average of the raw results for the given location and the two neighboring locations. (At -5 and 5, the raw value for the non-existent outer neighbor is taken to be equal to the raw value of that location.)

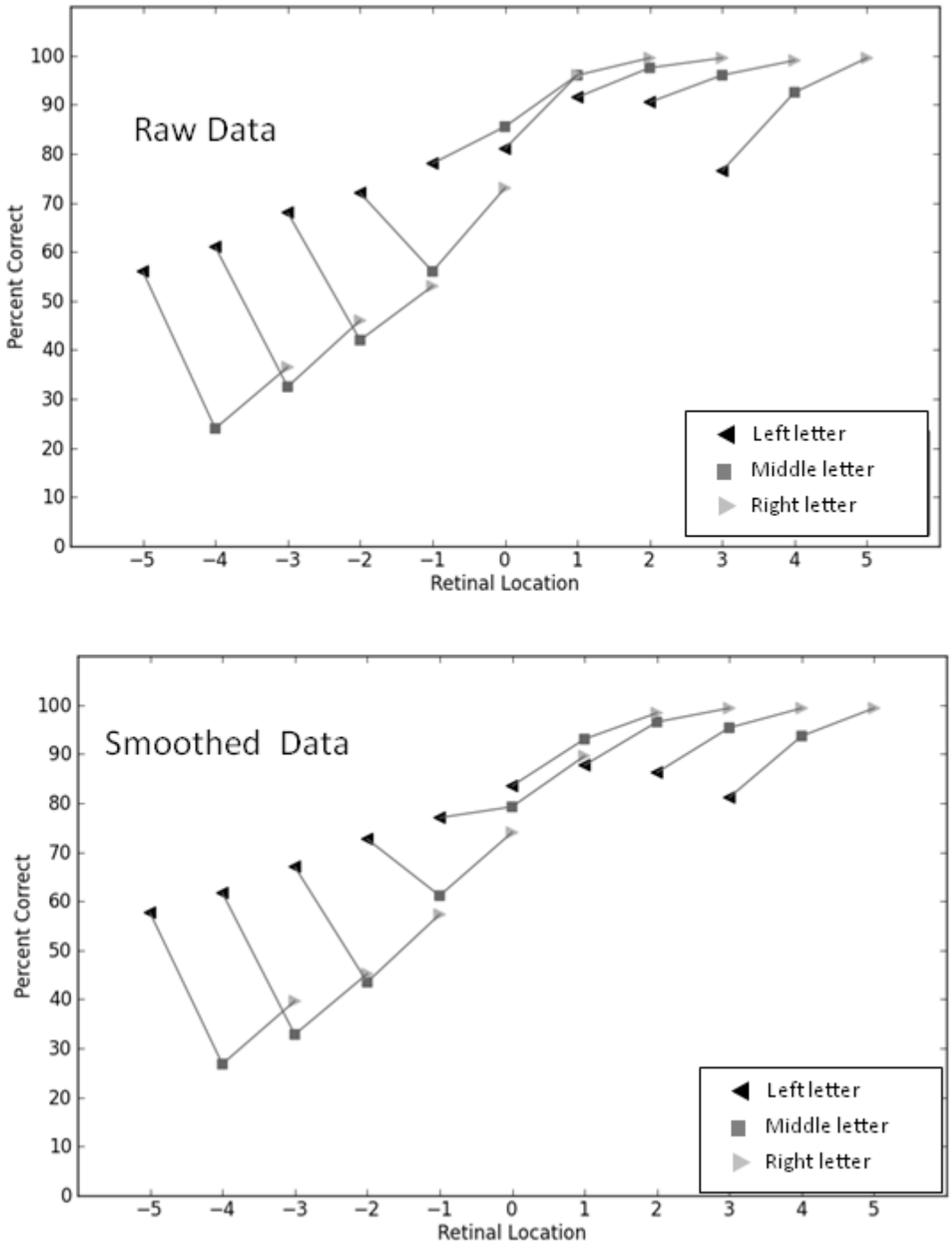

Figure 14: Trigram simulations for a RL language. Notation as in Figure 13.

Next we consider how the SERIOL2 mechanisms achieved the above goals. In the LVF, the interaction of the non-specific inhibition and the reduced Feature firing rate for inner letters caused a large effect of position for both reading directions; for the LR simulation, the additional influence of unidirectional inhibition yielded an even larger effect. (For the RL simulation, unidirectional inhibition was not a factor.) In the RVF, the reduced firing Feature firing rate for inner letters had a minimal effect (for both reading directions) because it was not magnified by non-specific lateral inhibition; for the RL simulation, unidirectional inhibition yielded a sizeable effect only at R_3. (For the LR simulation, unidirectional inhibition was not a factor.) Overall, the RVF positional effect is weaker than the LVF effect for both reading directions, satisfying the first goal.

For the RL simulation, $A(P_R) \approx A(P_M)$ at a given LVF location because both positions undergo the same reduced Feature firing rate and non-specific lateral inhibition. For the LR simulation, $A(P_R) < A(P_M)$ due to the additional effect of unidirectional inhibition. Hence, the simulations also satisfy the second goal.

The third and fourth goals are achieved via the weight gradients. For a LR language, L4→L3 weights are high for $R_{n\leq 0}$, yielding ceiling-level accuracy for P_L at these locations. The decreasing weight gradient yields decreasing accuracy for the RVF locations. For a RL language, the L4→L3 weights are high at $R_{n\geq 3}$, yielding ceiling-level accuracy for P_R at these locations. Weights decrease to the left, yielding reduced accuracy at central and LVF locations, where the non-specific inhibition accentuates the effect of the lower weights.

8.0 General Discussion

We presented experimental results on trigram identification which precisely matched the predictions of the SERIOL model for languages read from left to right, but not for languages read from right to left. As a result, the lower levels of the SERIOL model were modified to incorporate serial firing of Retinotopic-Letter representations, becoming the SERIOL2 model. We presented simulations illustrating how the SERIOL2 model explains the trigram data.

The goal of the simulations was to replicate the most salient aspects of the trigram data under a minimal number of assumptions. The key underlying assumptions are: (1) L2/3 RLetters tuned to

the LVF are located near each other in the LH. (2) The weight gradient on connections into L3 RLetters has the form shown in Figure 9. (3) Learned unidirectional inhibition originates from locations where the weight gradient is not decreasing. (4) RLetters that are cortically near each other non-specifically inhibit one another. (5) The outermost letter of a unilateral string has an attentional advantage.

We have sketched, in Section 5.4, how assumption (1) contributes to assumption (2). Assumptions (2) and (3) yield reading-direction-specific patterns of connectivity between RLetters. Assumptions (1) and (4) yield the pattern of non-specific inhibition. The resulting profiles of weight-gradient shapes, unidirectional inhibition, and non-specific inhibition were implemented in the normal simulations to yield serial firing across the L2 RLetters. The trigram simulations additionally implemented assumption (5), and the simulated accuracy patterns re produced the major characteristics of experimental results.

8.1 Lexical-level asymmetries

Next we consider implications of SERIOL2 for VF asymmetries observed at the lexical level, in LR languages. In the top graph of Figure 12, note that the slope increases as the location of the initial letter moves to the left, due to the unidirectional inhibition. For example, when the string starts at R_0, the final L2 RLetter fires ~30 ms after the first; when the string starts at R_{-5}, the final L2 RLetter fires ~60 ms after the first. The first L2 RLetter also starts firing later when it is located further from fixation. Hence, by ~80 ms, all six L2 RLetters have fired for a string starting at R_0, while only the first three L2 RLetters have fired for a string starting at R_{-5}.

This difference explains recent masked-priming results in a study which employed lexical decision on six-letter stimuli (Van der Haegen, Brysbaert, & Davis, 2009). The prime was formed by transposing or replacing the target's second / third letters, or fourth / fifth letters. For example, consider the target "CARPET"; transposition primes would be "crapet" or "carept", and replacement primes could be "cumpet" or "carmot". Fixation position was also varied. The prime and target were presented such that fixation fell at all possible between-letter positions; that is, fixation fell between letter *n* and *n* +1, for *n* = 1,2,3,4,5. (Within a trial, prime and target were both presented with the same fixation position). We concentrate on the conditions where the prime/target fell mostly in the RVF (i.e., fixation between first and second letters) or the LVF (i.e., fixation between fifth and sixth letters).

A prime with transposed letters is facilitative compared to a prime with replaced letters (Perea & Lupker, 2003), presumably because the transposed letters contribute to the activation of target Open-Bigrams, while replaced letters do not (Grainger & Whitney, 2004). Hence, a transposition prime should yield faster reaction times than a replacement prime only if the corresponding ALetters have been activated. For example, the prime "carept" should yield faster reaction times than "carmot" only if the ALetters representing the fourth and fifth letters have been activated. We consider observed transposition advantages – the decrease in reaction time for transposed versus replaced primes.

Let $T(p, V)$ denote the transposition advantage for prime manipulation at positions p and $p+1$, for stimuli falling mostly in visual field *V*. For example, $T(4,$ LVF) denotes the difference in reaction times for a prime like "carept" versus a prime like "carmot", when the prime/target is fixated between the fifth and sixth letters.

The experimental results were as follows: $T(2, \text{RVF}) = -28$ ms, $T(4, \text{RVF}) = -39$ ms, $T(2, \text{LVF}) = -62$ ms, and $T(4, \text{LVF}) = 1$ ms. That is, transposition advantages occurred for all conditions, except the fourth/fifth position with LVF stimuli. The LVF results provide clear evidence of serial processing, indicating that the prime's second and third letters, but not its fourth and fifth letters, successfully activated the corresponding ALetters. In contrast, the prime's first through fifth letters activated the corresponding ALetters when the prime occurred mostly in the RVF, consistent with the SERIOL2 proposal of faster L2 RLetter and (and consequently ALetter) activation for RVF stimuli.

The slower rate of ALetter activation for LVF stimuli explains other VF asymmetries in lexical decision. Reaction times for LVF presentation are longer than for RVF presentation. For RVF presentation, word length and orthographic-neighborhood density have no effect on reaction times; for LVF presentation, increased length is inhibitory, while increased neighborhood density is facilitative (Lavidor & Ellis, 2002). In the following, we use the term *Target Bigram* to denote an Open-Bigram with an excitatory connection to the VWF encoding the stimulus word. Recall that, in SERIOL2, each open-bigram is represented by a pool of Open-Bigrams having different temporal sensitivities; the total number of activated Open-Bigrams unit in each pool decreases as the interval between the firing of the constituent ALetters increases. The increased time between the firing of consecutive ALetters under LVF presentation will result in failure to activate some

of the Target Bigrams. We denote Target Bigrams that were not activated by ALetters as *quiescent Target Bigrams*. LVF presentation activates fewer Target Bigrams than RVF presentation, yielding less excitation of the target VWF. This explains why reaction times are longer for LVF stimuli.

Next we consider the effect of the feedback connections from VWFs to Open-Bigrams. Convergent feedback from VWFs can activate a quiescent Target Bigram. Activation of quiescent Target Bigrams would increase input to the target VWF, causing faster settling at the VWF level and decreased reaction times. Higher neighborhood density would increase the number of activated VWFs, which would increase the probability of activation of quiescent Target Bigrams. Hence, higher neighborhood density is facilitative under LVF presentation because it increases the probability of activation of quiescent Target Bigrams. Longer words generally have lower neighborhood densities than shorter words, so longer words are less likely to activate quiescent Target Bigrams than shorter words. Hence, increased word length is inhibitory under LVF presentation. In contrast, under RVF presentation, all Target Bigrams are activated by ALetter stimulation, yielding no quiescent Target Bigrams. Therefore top-down excitation cannot activate additional Target Bigrams, and neighborhood density has no effect on reaction times.

Whitney (2011) demonstrated that these assumptions can indeed account for VF-specific patterns in the effects of neighborhood density and word length, via a large-scale spiking-neuron simulation of the Open-Bigram and VWF levels.[1] Although the simulation utilized serial activation of Open-Bigrams, settling time at the VWF level did *not* show a length effect under simulated central or RVF presentation. However, settling time showed a length effect under simulated LVF presentation. This LVF length effect was due to decreased top-down activation of quiescent Target Bigrams for longer words, not directly to seriality, as the length effect disappeared when neighborhood density was matched across word length.

[1] The account presented in Whitney (2011) of why the interval between successive ALetters is longer in the LVF than the RVF was based on a preliminary version of the SERIOL2 model. The present specification of SERIOL2 supersedes the one sketched in Whitney (2011).

8.2 Comparison of SERIOL2 to Related Models

The only other model to address the issue of hemisphere-specific orthographic processing is the the Modified Receptive Field (MRF) model (Chanceaux & Grainger, 2012). For a left-to-right language, the MRF posits that crowding for an LVF letter is stronger from neighboring letters to the left than to the right, whereas crowding for an RVF letter is roughly similar for neighboring letters to the left and right. The MRF specifies that this VF asymmetry in crowding instantiates a specialization for detection of the initial letter of a word, which usually falls in the LVF in a left-to-right language.

The MRF recapitulates a prediction that was inherent to the SERIOL model (Whitney, 2001). In contrast to the MRF, the VF asymmetry in SERIOL is directly related to letter-position encoding. The MRF model (like SERIOL) predicts that, in Hebrew, RVF letters should experience more crowding from neighboring letters to the right than to the left. However, we have seen that this prediction is incorrect, as there was no advantage for an initial versus a second letter within R_1, R_2 or R_3 in Hebrew. As a result, the retinotopic mechanisms driving the serial letter encoding in SERIOL were re-specified, yielding SERIOL2. It is unclear how the MRF could be modified to become consistent with the Hebrew data, as the MRF specifies that the VF asymmetry exists only to provide an advantage for an initial letter in the VF in which it usually occurs; however, no such advantage is observed in Hebrew.

The MRF model is an extension of the Parallel Open-Bigram (POB) model (Grainger & van Heuven, 2003), which posits that Retinotopic-Letters activate (non-retinotopic) Open-Bigrams in parallel. However, the POB model does not specify how this is accomplished. We suggest that it would require a layer of Retinotopic Open-Bigrams between the Retinotopic-Letters and (non-retinotopic) Open-Bigrams. Note that that the POB model does not include an abstract (non-retinotopic) representation at the level of individual letters. It is generally acknowledged that Open-Bigrams do not provide a suitable encoding for processing along the dorsal phonological pathway. Hence the POB model does not address how an abstract representation of letter order is encoded for phonological processing. In contrast, the serial letter encoding in the SERIOL/SERIOL2 model provides such an abstract representation, which can then be parsed into a graphemic representation on the dorsal pathway.

The Local Combination Detector model (LCD) (Dehaene et al., 2005) proposes parallel activation of noisy Retinotopic Letter, Bigram, and Quadrigram representations. Like the POB model, the LCD model does not address the issue of how letter order is encoded for the dorsal pathway.

The only other model to include a serial letter encoding is the Spatial Coding model (Davis, 2010). The Spatial Coding model specifies that an activation gradient across Retinotopic-Letters is converted into a phase (i.e., serial) encoding of letter order, via an unspecified scanning process. On the ventral pathway, this serial encoding activates VWFs by "superposition matching", which entails complex computations local to every VWF. That is, each VWF has its own set of letter representations, and the serial letter representation of the stimulus interacts with each set of VWF letter representations to yield varying VWF activations. We note that the simulation of the Spatial Coding model presented in Davis (2010) did not actually implement the superposition matching process within the neural network. Rather, OWF activation levels were simply computed according to a formula.

In contrast to the Spatial Coding model, SERIOL2 specifies precisely how the serial ALetter encoding is induced. ALetters then activate a single set of Open-Bigrams, which connects to all VWFs; Open-Bigram → VWF activation has been simulated in a spiking neural network (Whitney, 2011). Hence, the specification of how the letter-level encoding activates OWFs is much simpler in SERIOL2 than in the Spatial Coding model, and the SERIOL2 mechanism (of Open-Bigrams) has been implemented within a neural-network framework.

Other accounts of orthographic processing (Overlap model: Gomez et al., 2008; a non-model: Norris & Kinoshita, 2012) do not address the issue of how a retinotopic representation is converted into an abstract representation of letter order.

We have seen that VF-specific patterns are quite different for Hebrew versus English trigram identification, and that these patterns are highly robust across subjects within a reading direction. Therefore, any account of orthographic processing should explain these patterns.

8.3 Future Research

Future theoretical and modeling work will focus on the issue of how the SERIOL2 architecture arises during reading acquisition. Briefly, we conjecture that the left lateralization of

orthographic processing originates from top-down influences generated by learned Grapheme-Phoneme mappings, which causes the instantiation of ALetters in left IOG. Competitive, associative, and perceptual learning, driven by attentional patterns during the early phases of reading, support the formation of RLetters and their proposed connectivity. Serial letter processing progresses from explicit letter-by-letter decoding via sequential fixations, to seriality driven by top-down attentional signals within a fixation, to a serial encoding induced in a bottom-up manner via the weight gradients and unidirectional inhibition specified in the SERIOL2 model.

An fMRI experiment could directly test the SERIOL2 proposal that LVF letters are represented near the foveal center in left V4. Because the area of foveal V4 activated by LVF strings is predicted to be small and the location of foveal V4 could vary across subjects, analyses within individual subjects would be necessary to directly detect the predicted activity. Retinotopic mapping could be used to locate the representation of the V4 fovea and parafovea, as the retinotopic visual areas can be separated even within the foveal confluence (Schira, Tyler, Breakspear, & Spehar, 2009). LVF presentation of letter strings should activate left foveal V4, while RVF presentation should not activate right foveal V4.

Another important direction for future research is to perform the trigram experiments with other subject groups. The adult skilled readers who undertook the present experiments showed strong, distinctive asymmetries across VFs in both English and Hebrew. We take these patterns to be a signature of specialized orthographic processing. The consistency of the results across subjects indicates that the trigram protocol could be used with children to evaluate the normal time course of the acquisition of orthographic processing, and to detect individuals with abnormal orthographic analysis. We suggest that at least two types of aberrant trigram patterns may be identified.

One possible aberrant trigram pattern is a symmetric effect of position across VFs, with little effect of position in either VF. Close examination of the trigram data presented in Figure 6 in a study of a seventh-grade dyslexic and seven age-matched controls (Dubois et al., 2007) reveals that the dyslexic participant showed little effect of position in the LVF, while all of the controls showed a strong effect. The dyslexic's symmetric pattern indicates failure to acquire specialized orthographic processing. This condition may stem from inability to form automatic Grapheme-

Phoneme mappings, which would prevent the cascade to left-lateralized letter processing (Blomert, 2011). A deficit in the capacity to narrow attention to a single letter during the stage of letter-by-letter decoding may limit the formation of automatic Grapheme-Phoneme mappings (Franceschini, Gori, Ruffino, Pedrolli, & Facoetti, 2012).

Another possible trigram pattern is normal VF asymmetry with selectively reduced perception of the middle letter, which we have observed in college students with dyslexia (Callens, Whitney, Tops, & Brysbaert, in press). For the dyslexics, but not the control subjects, middle-letter accuracy correlated with speeded word-reading ability. This dyslexic profile reflects increased crowding between letters, with normal left-lateralization of orthographic processing. To understand possible origins of this deficit, future modeling research will also focus on the process of letter recognition.

Hence, different abnormal patterns of trigram identification may signal different underlying deficits preventing the acquisition of skilled orthographic processing. We suggest that more precise characterization of trigram patterns and underlying deficits could lead to methods of reading remediation that are specifically targeted to individual subjects.

In conclusion, the trigram results suggested that the SERIOL mechanisms originally proposed for instantiation of the locational gradient (i.e., differential bottom-up weights, and unidirectional lateral inhibition) instead directly induce serial firing across retinotopic letter representations, forming the basis of the SERIOL2 model. The model and the trigram protocol offer new directions for experimental research in the quest to understand skilled orthographic processing, and the trajectory of its successful or unsuccessful acquisition.

Appendix

The simulations were implemented using the Brian package (Goodman & Brette, 2009). The simulations employed leaky integrate-and-fire neurons, whose membrane potential V is governed by the following ordinary differential equation:

$$\tau_m \frac{dV}{dt} = -(V - V_R) + I(t)$$

where τ_m is the membrane time constant, V_R is the resting potential, and $I(t)$ is the input current generated by incoming spikes. V_R was set to 0 mV, for convenience. The input current is given by:

$$I(t) = \sum_{i,n} w_i * \Theta\left(t - \overline{t_i^n}\right) * \exp\left(-\frac{t - \overline{t_i^n}}{\tau_s}\right)$$

where $\boldsymbol{w_i}$ denotes the weight on synapse $\boldsymbol{i}$, $\boldsymbol{\Theta}$ is the Heaviside function, $\overline{\boldsymbol{t_i^n}}$ denotes the arrival time of the *n*th spike at synapse $\boldsymbol{i}$, and $\boldsymbol{\tau_s}$ is the synaptic time constant. (Due to axonal delay, the arrival time of generated spike can be $> t$, so the Heaviside function is used to restrict arrival times to $\leq \boldsymbol{t}$.) Whenever $\boldsymbol{V}$ reaches the spiking threshold, $\boldsymbol{T}$, a spike is emitted and $\boldsymbol{V}$ is reset to $\boldsymbol{V_R}$. The differential equation is solved numerically (Euler method) with a time step of 0.1 ms.

The values used for the weight gradients are given in Table A.1. The weight gradient is multiplied by $0.2T$ to yield the weights for L4→L3 RLetter connections. Values ranging from $-0.03T$ to $-0.24T$ were used for the lateral inhibitory weights on L3→L2 RLetter connections. Other parameters are given in Table A.2. Some parameters are unrealistic for single neurons, such as the Feature firing rate; such parameters should be interpreted as the net effect of a group of neurons. In the trigram simulations, the mask time M_i for the *i*th run of a given trigram location is given by:

$$M_i = M_{\min} + \frac{i * (M_{\max} - M_{\min})}{N}$$

where $M_{\min} = 30$ ms, $M_{\max} = 70$ ms, and $N = 200$ is the total number of runs per location.

Ecc.	-5	-4	-3	-2	-1	0	1	2	3	4	5
LR	0.80	0.83	0.86	0.90	0.95	1.00	0.80	0.64	0.51	0.43	0.38
RL	0.40	0.42	0.44	0.47	0.50	0.54	0.64	0.80	1.00	0.90	0.80

Table A.1 : Parameter values for the weight gradients, where **Ecc.** denotes eccentricity and **LR** and **RL** denote reading direction.

Parameter Name	Value	Description
nLocs	11	Number of neurons per layer
featRate	5000 Hz	Feature firing rate
maskRate	35 Hz	Mask firing rate, for trigrams
decayDur	40 ms	Time for Feature firing rate to decay to 0 Hz after mask for trigrams
reduFac	70%	Reduced Feature firing rate for a non-fixated trigram
reduFacF	82%	Reduced Feature firing rate for a fixated trigram
T	10 mv	Spiking threshold
eWtF4	1.5*T*	Weight for Features →L4
eWt33	1.5*T*	Weight for L3 self-excitation
eWt32	0.7*T*	Weight for L3 → L2
iWt33	-0.1*T*	Weight for non-specific inhibition (L3 → L3)
iWt32	-10*T*	Weight for feedback Inhibition (L2 → L3)
iWtM	-0.7*T*	Mask weight, for trigrams
tauVp	10 ms	Membrane time constant
tauE	50 ms	EPSP synaptic time constant
tauFI	15 ms	Fast IPSP synaptic time constant (non-specific inhibition)
tauSI	50 ms	Slow IPSP synaptic time constant (unidirectional inhibition)
tauRI	200 ms	Feedback IPSP time constant
seDelay	2 ms	Delay in self-excitation for L3

Table A.2: Other parameter values used in the simulations. Parameter names are as in the Brian code.

Acknowledgements

We thank A. Udaya Shankar for valuable comments.

References

Adelman, J. S. (2011). Letters in time and retinotopic space. *Psychological Review*, *118*(4), 570–582. doi:10.1037/a0024811

Adelman, J. S., Marquis, S. J., & Sabatos-DeVito, M. G. (2010). Letters in words are read simultaneously, not in left-to-right sequence. *Psychological science*, *21*(12), 1799–1801. doi:10.1177/0956797610387442

Barca, L., Cornelissen, P., Simpson, M., Urooj, U., Woods, W., & Ellis, A. W. (2011). The neural basis of the right visual field advantage in reading: an MEG analysis using virtual electrodes. *Brain and Language*, *118*(3), 53–71. doi:10.1016/j.bandl.2010.09.003

Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K., & Wandell, B. A. (2007). Differential sensitivity to words and shapes in ventral occipito-temporal cortex. *Cerebral cortex (New York, N.Y.: 1991)*, *17*(7), 1604–1611. doi:10.1093/cercor/bhl071

Besserve, M., Philippe, M., Florence, G., Laurent, F., Garnero, L., & Martinerie, J. (2008). Prediction of performance level during a cognitive task from ongoing EEG oscillatory activities. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*, *119*(4), 897–908. doi:10.1016/j.clinph.2007.12.003

Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., & Buchanan, L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *NeuroImage*, *33*(2), 739–748. doi:10.1016/j.neuroimage.2006.06.053

Blomert, L. (2011). The neural signature of orthographic–phonological binding in successful and failing reading development. *NeuroImage*, *57*(3), 695–703. doi:10.1016/j.neuroimage.2010.11.003

Bosse, M.-L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: the visual attention span deficit hypothesis. *Cognition*, *104*(2), 198–230. doi:10.1016/j.cognition.2006.05.009

Callens, M., Whitney, C., Tops, W., & Brysbaert, M. (in press). No deficiency in left-to-right processing of words in dyslexia but evidence for enhanced visual crowding. *Quarterly Journal of Experimental Psychology*.

Carreiras, M., Duñabeitia, J. A., & Molinaro, N. (2009). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. *Cerebral Cortex (New York, N.Y.: 1991)*, *19*(11), 2659–2670. doi:10.1093/cercor/bhp019

Chanceaux, M., & Grainger, J. (2012). Serial position effects in the identification of letters, digits, symbols, and shapes in peripheral vision. *Acta psychologica*, *141*(2), 149–158. doi:10.1016/j.actpsy.2012.08.001

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: the case for the visual word form area. *NeuroImage*, *22*(1), 466–476. doi:10.1016/j.neuroimage.2003.12.049

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., & Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain: a journal of neurology*, *123 (Pt 2)*, 291–307.

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713–758. doi:10.1037/a0019738

Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding: evidence from orthographic similarity effects. *Journal of Experimental Psychology. Human Perception and Performance*, *32*(3), 535–557. doi:10.1037/0096-1523.32.3.535

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences*, *9*(7), 335–341. doi:10.1016/j.tics.2005.05.004

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J.-B., Le Bihan, D., & Cohen, L. (2004). Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence. *Psychological Science: A Journal of the American Psychological Society / APS*, *15*(5), 307–313. doi:10.1111/j.0956-7976.2004.00674.x

Douglas, R. J., & Martin, K. A. C. (2004). Neuronal circuits of the neocortex. *Annual review of neuroscience*, *27*, 419–451. doi:10.1146/annurev.neuro.27.070203.144152

Dubois, M., De Micheaux, P. L., Noël, M.-P., & Valdois, S. (2007). Preorthographical constraints on visual word recognition: evidence from a case study of developmental surface dyslexia. *Cognitive Neuropsychology*, *24*(6), 623–660. doi:10.1080/02643290701617330

Fiebach, C. J., Friederici, A. D., Müller, K., & von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience*, *14*(1), 11–23. doi:10.1162/089892902317205285

Fischer-Baum, S., Charny, J., & McCloskey, M. (2011). Both-edges representation of letter position in reading. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-011-0160-3

Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., & Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Current biology: CB*, *22*(9), 814–819. doi:10.1016/j.cub.2012.03.013

Frankish, C., & Barnes, L. (2008). Lexical and sublexical processes in the perception of transposed-letter anagrams. *Quarterly Journal of Experimental Psychology (2006)*, *61*(3), 381–391. doi:10.1080/17470210701664880

Frost, R. (2012). Towards a universal model of reading. *The Behavioral and brain sciences*, 1–17. doi:10.1017/S0140525X11001841

Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the "visual word form area." *Neuron*, *62*(2), 199–204. doi:10.1016/j.neuron.2009.03.017

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: a model of letter position coding. *Psychological Review*, *115*(3), 577–600. doi:10.1037/a0012667

Goodman, D. F. M., & Brette, R. (2009). The Brian simulator. *Frontiers in Neuroscience*, *3*(2), 192–197. doi:10.3389/neuro.01.026.2009

Grainger, J., Granier, J.-P., Farioli, F., Van Assche, E., & van Heuven, W. J. B. (2006). Letter position information and printed word perception: the relative-position priming constraint. *Journal of Experimental Psychology. Human Perception and Performance*, *32*(4), 865–884. doi:10.1037/0096-1523.32.4.865

Grainger, J., Tydgat, I., & Isselé, J. (2010). Crowding affects letters and symbols differently. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(3), 673–688. doi:10.1037/a0016888

Grainger, J., & van Heuven, W. (2003). Modeling letter position coding in printed word perception. In *Bonin, P, ed. Mental lexicon: Some words to talk about words* (pp. 1–23). New York: Nova Science.

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences*, *8*(2), 58–59. doi:10.1016/j.tics.2003.11.006

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 (New York, N.Y.: 1991)*, *9*(5), 476–483.

Holcomb, P. J., & Grainger, J. (2006). On the time course of visual word recognition: an event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience*, *18*(10), 1631–1643. doi:10.1162/jocn.2006.18.10.1631

Jordan, T. R., Fuggetta, G., Paterson, K. B., Kurtev, S., & Xu, M. (2011). An ERP Assessment of Hemispheric Projections in Foveal and Extrafoveal Word Recognition. *PloS One*, *6*(9), e23957. doi:10.1371/journal.pone.0023957

Keysers, C., & Perrett, D. I. (2002). Visual masking and RSVP reveal neural competition. *Trends in Cognitive Sciences*, *6*(3), 120–125. doi:16/S1364-6613(00)01852-0

Kwon, M., Legge, G. E., & Dubbels, B. R. (2007). Developmental changes in the visual span for reading. *Vision Research*, *47*(22), 2889–2900. doi:10.1016/j.visres.2007.08.002

Lavidor, M., & Ellis, A. W. (2002). Word Length and Orthographic Neighborhood Size Effects in the Left and Right Cerebral Hemispheres. *Brain and Language*, *80*(1), 45–62. doi:10.1006/brln.2001.2583

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*(6), 725–743.

Lisman, J. E., & Idiart, M. A. (1995). Storage of 7 +/- 2 short-term memories in oscillatory subcycles. *Science (New York, N.Y.)*, *267*(5203), 1512–1515.

Lübke, J., Roth, A., Feldmeyer, D., & Sakmann, B. (2003). Morphometric Analysis of the Columnar Innervation Domain of Neurons Connecting Layer 4 and Layer 2/3 of Juvenile Rat Barrel Cortex. *Cerebral Cortex*, *13*(10), 1051 –1063. doi:10.1093/cercor/13.10.1051

Mano, Q. R., Humphries, C., Desai, R. H., Seidenberg, M. S., Osmon, D. C., Stengel, B. C., & Binder, J. R. (2012). The Role of Left Occipitotemporal Cortex in Reading: Reconciling Stimulus, Task, and Lexicality Effects. *Cerebral cortex (New York, N.Y.: 1991)*. doi:10.1093/cercor/bhs093

Mariño, J., Schummers, J., Lyon, D. C., Schwabe, L., Beck, O., Wiesing, P., … Sur, M. (2005). Invariant computations in local cortical networks with balanced excitation and inhibition. *Nature neuroscience*, *8*(2), 194–201. doi:10.1038/nn1391

Martin, C. D., Thierry, G., Démonet, J.-F., Roberts, M., & Nazir, T. (2007). ERP evidence for the split fovea theory. *Brain research*, *1185*, 212–220. doi:10.1016/j.brainres.2007.09.049

Nestor, A., Behrmann, M., & Plaut, D. C. (2012). The Neural Basis of Visual Word Form Processing: A Multivariate Investigation. *Cerebral cortex (New York, N.Y.: 1991)*. doi:10.1093/cercor/bhs158

Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: why there's nothing special about the perception of orthography. *Psychological review*, *119*(3), 517–545. doi:10.1037/a0028450

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition*, *31*(6), 829–841.

Petrov, Y., & Meleshkevich, O. (2011). Locus of spatial attention determines inward-outward anisotropy in crowding. *Journal of vision*, *11*(4). doi:10.1167/11.4.1

Petrov, Y., Popple, A. V., & McKee, S. P. (2007). Crowding and surround suppression: not to be confused. *Journal of vision*, *7*(2), 12.1–9. doi:10.1167/7.2.12

Rauschecker, A. M., Bowen, R. F., Parvizi, J., & Wandell, B. A. (2012). Position sensitivity in the visual word form area. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(24), E1568–1577. doi:10.1073/pnas.1121304109

Richardson, F. M., Seghier, M. L., Leff, A. P., Thomas, M. S. C., & Price, C. J. (2011). Multiple routes from occipital to temporal cortices during reading. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *31*(22), 8239–8247. doi:10.1523/JNEUROSCI.6519-10.2011

Schira, M. M., Tyler, C. W., Breakspear, M., & Spehar, B. (2009). The foveal confluence in human visual cortex. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *29*(28), 9050–9058. doi:10.1523/JNEUROSCI.1760-09.2009

Seghier, M. L., Neufeld, N. H., Zeidman, P., Leff, A. P., Mechelli, A., Nagendran, A., … Price, C. J. (2012). Reading without the left ventral occipito-temporal cortex. *Neuropsychologia*. doi:10.1016/j.neuropsychologia.2012.09.030

Siegel, M., Warden, M. R., & Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(50), 21341–21346. doi:10.1073/pnas.0908193106

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., & Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *NeuroImage*, *56*(1), 330–344. doi:16/j.neuroimage.2011.01.073

Thesen, T., McDonald, C. R., Carlson, C., Doyle, W., Cash, S., Sherfey, J., … Halgren, E. (2012). Sequential then interactive processing of letters and words in the left fusiform gyrus. *Nature communications*, *3*, 1284. doi:10.1038/ncomms2220

Van den Broeck, W., & Geudens, A. (2012). Old and new ways to study characteristics of reading disability: The case of the nonword-reading deficit. *Cognitive psychology*, *65*(3), 414–456. doi:10.1016/j.cogpsych.2012.06.003

Van der Haegen, L., Brysbaert, M., & Davis, C. J. (2009). How does interhemispheric communication in visual word recognition work? Deciding between early and late integration accounts of the split fovea theory. *Brain and Language*, *108*(2), 112–121. doi:10.1016/j.bandl.2008.06.005

Van der Haegen, L., Drieghe, D., & Brysbaert, M. (2010). The Split Fovea Theory and the Leicester critique: what do the data say? *Neuropsychologia*, *48*(1), 96–106. doi:10.1016/j.neuropsychologia.2009.08.014

Van der Mark, S., Bucher, K., Maurer, U., Schulz, E., Brem, S., Buckelmüller, J., … Brandeis, D. (2009). Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *NeuroImage*, *47*(4), 1940–1949. doi:10.1016/j.neuroimage.2009.05.021

Van Essen, D. C., Newsome, W. T., & Bixby, J. L. (1982). The pattern of interhemispheric connections and its relationship to extrastriate visual areas in the macaque monkey. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *2*(3), 265–283.

Velan, H., & Frost, R. (2009). Letter-transposition effects are not universal: The impact of transposing letters in Hebrew. *Journal of Memory and Language*, *61*(3), 285–302. doi:10.1016/j.jml.2009.05.003

Vergara-Martínez, M., & Swaab, T. Y. (2012). Orthographic neighborhood effects as a function of word frequency: an event-related potential study. *Psychophysiology*, *49*(9), 1277–1289. doi:10.1111/j.1469-8986.2012.01410.x

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: the SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, *8*(2), 221–243.

Whitney, C. (2011). Location, location, location: how it affects the neighborhood (effect). *Brain and Language*, *118*(3), 90–104. doi:10.1016/j.bandl.2011.03.001

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*(3), 274–301.

Wilson, S. M., Brambati, S. M., Henry, R. G., Handwerker, D. A., Agosta, F., Miller, B. L., … Gorno-Tempini, M. L. (2009). The neural basis of surface dyslexia in semantic dementia. *Brain: A Journal of Neurology*, *132*(Pt 1), 71–86. doi:10.1093/brain/awn300

Wilson, T. W., Leuthold, A. C., Moran, J. E., Pardo, P. J., Lewis, S. M., & Georgopoulos, A. P. (2007). Reading in a deep orthography: neuromagnetic evidence for dual-mechanisms. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *180*(2), 247–262. doi:10.1007/s00221-007-0852-0