Does Letter-by-Letter Reading in Pure Alexia Reflect Seriality in Normal Reading?

Carol Whitney

Abstract

Left occipitotemporal damage often yields pure alexia, which is characterized by an extreme length effect in reading single words. Based on the SERIOL model, I propose that the overt seriality displayed by pure alexics is necessary to replace the very rapid, automatic seriality normally produced by left occipitotemporal cortex. In supporting this account, I extend the SERIOL model to include the phonological route, make specific proposals as to the nature of hemispheric specializations in visual processing, and discuss the relationship between string processing and visual object recognition.

Introduction

In the unfortunate event of left occipitotemporal damage, the result is often pure alexia, which is characterized by slow, accurate reading with spared spelling ability. The hallmark of this condition is an extreme length effect, where each additional letter can increase reaction times by 0.5 to 3 *seconds*. While the per-letter processing time may decrease during recovery, a strong length effect persists over time (Beeson, Magloire & Robey, 2005). In contrast, normal readers show no length effect for words (Weekes, 1997). Thus another name for pure alexia is letter-by-letter (LBL) reading, and eye movement studies have confirmed that pure alexic patients do indeed read letter by letter (Behrmann, Shomstein, Black & Barton, 2001; Johnson & Rayner, 2007).

In skilled readers, a region of left occipitotemporal cortex, dubbed the Visual Word Form Area, becomes tuned to one's native orthography (Baker et al., 2007; McCandliss, Cohen & Dehaene, 2003). Pure alexia generally arises via damage to, or disconnection from, this area, and the homologous right occipitotemporal region likely supports LBL reading (Cohen et al., 2003; 2004).

Thus, it would appear that the left hemisphere (LH) supports fast, parallel lexical access while the right hemisphere (RH) can only provide slow, serial lexical access. But why is this the case? Why should the RH be incapable of processing letters in parallel? This is especially puzzling given that the RH is generally recognized as being specialized for visuospatial analysis, which entails parallel processing.

This article proposes a different way of looking at the puzzle, namely that RH parallel processing is actually the *cause* of LBL reading. Briefly, I propose that the LH normally provides extremely rapid *serial* processing of letters and that lexical access requires a serial encoding of letter order. When the LH is damaged, the only way to provide the necessary serial input is via explicit letter-by-letter reading.

In the remainder of this article, I flesh out and support this proposal, which stems from the SERIOL model of orthographic encoding (Whitney, 2001; 2004). While some aspects of this proposal are admittedly quite speculative, it does offer a cohesive account of letter processing, hemispheric specialization, and pure alexia. In the following, I first review the SERIOL model. I then make novel suggestions as to the nature of hemispheric specializations in visual analysis. Next, I extend the SERIOL model to include the phonological route, and then present the account of pure alexia. I conclude with a discussion establishing that absence of a length effect in normal visual word recognition does not preclude the proposed serial encoding.

Review of the SERIOL Model

It is generally recognized that there are two reading routes. On the occipitotemporal (ventral) route, orthography is mapped onto a lexicosemantic representation. On the occipitoparietal (dorsal) route, orthography is mapped to phonological encoding. The original SERIOL model, described next, addressed processing on the ventral route. For brevity, the following review is limited to aspects of the model relevant to the present discussion. For a more detailed specification of the model, including motivations for the proposed representations, see Whitney (2001; 2004).

The model is a verbal one that specifies how a retinotopic representation is transformed into an abstract encoding of letter order, and how this encoding then activates lexical representations. It is comprised of five layers: edges, features, letters, open-bigrams, and words.

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

The feature layer models ventral retinotopic extrastriate areas (hV4, VO1, VO2). The model proposes the acuity gradient is transformed into a decreasing activation gradient at the feature layer, dubbed the *locational gradient*. That is, activation level is highest for the first letter, and decreases across the string. (Note that although this is called the "feature" layer, whole-letter representations are not ruled out; the important point is the proposal of an activation gradient across a *retinotopic* encoding.) The formation of the locational gradient is based on hemisphere-specific processing that is learned during reading acquisition in response to a top-down attention gradient (Whitney & Cornelissen, 2005). Hemisphere-specific processing is required because the acuity gradient matches the locational gradient in one visual field (VF) / hemisphere, but not the other. For example, in language read from left to right, the locational gradient *decreases* from left to right (in both the LVF/RH and RVF/LH), while the acuity gradient *increases* from left to right in the LVF/RH (i.e., from the periphery toward fixation), but *decreases* in the RVF/LH (i.e., from fixation toward the periphery). Thus the acuity and locational gradients do *not* match in the LVF/RH, while they do match in RVF/LH. Therefore, more extensive processing is required in the RH in order to convert the increasing acuity gradient into the decreasing locational gradient. For details of the proposed hemisphere-specific processing, see Whitney (2001).

The letter layer corresponds to left posterior/middle fusiform gyrus. Here letter units fire serially. That is, the letter unit encoding the first letter fires, then the unit encoding the second letter fires, etc. This mechanism is based on the general proposal that a sequence is encoded in successive gamma cycles (60 Hz) of a theta cycle (5 Hz) (Lisman & Idiart, 1995). Thus, each letter unit is active for about 15 ms. This firing pattern is induced by the interaction of the locational gradient with synchronous subthreshold oscillations within the letter units. The amount of input from the feature layer determines time of firing at the letter layer, with stronger input inducing earlier firing. As discussed below, this is the key point of abstraction, as the serial letter encoding is not retinotopic, yet it represents the spatial relationships between letters.

At the open-bigram (Grainger & Whitney, 2004) layer, corresponding to the Visual Word Form Area in left middle fusiform, units recognize pairs of letter units that fire in a particular order. That is, open-bigram unit **XY** is activated when letter-unit **X** fires before **Y** (where the letters *x* and *y* were not necessarily contiguous in the string). The open-bigram units activate the word layer, via weighted connections. The word layer comprises an encoding of orthographic word forms in left anterior fusiform.

In summary, the symmetric acuity gradient of the edge layer is converted into a monotonically decreasing activation gradient in the feature layer, called the locational gradient. The locational gradient interacts with oscillatory letter units to produce a serial encoding at the letter layer. The serial encoding activates open-bigram units, which recognize letter pairs that fire in a particular order. The open-bigrams then activate lexical representations in the word layer.

Note that the model is agnostic as to whether letters are identified serially, or are identified in parallel and then mapped to a serial encoding. That is, the retinotopic "features" in the feature layer could be letter fragments or entire letters. However, an asymmetric effect of non-letter flankers suggests that the locational gradient is established at a sub-letter level (Whitney, in press), indicating that the feature layer includes fragments. It is possible, though, that the feature layer is actually composed of different sublevels. For example, the locational gradient may initially be established over a sublevel that encodes fragments, and then further tuned over a sublevel that encodes case-specific letters. (Such parallel identification of letters would have the advantage that it solves the problem of how to bind fragments together.) This retinotopic letter sublevel would then activate the abstract letter layer. Thus the model makes the strong claim that a retinotopic activation gradient induces a serial encoding at the letter level, but does not make a strong claim about when and where letters are actually first identified. Detailed fMRI priming studies are planned to resolve this issue.

Hemispheric Asymmetries

Next we consider hemispheric asymmetries for visual processing in general. Although this may seem to take us rather far afield, the nature of hemispheric representations is crucial to the question of why the LH, but not RH, supports an orthographic encoding that allows rapid lexical access. This discussion will focus on how location invariance is achieved, and how relationships between object parts are represented.

It is well known that left occipitotemporal damage typically interferes with reading, while right occipitotemporal damage interferes with face processing. To account for this pattern, it has been proposed that the LH is specialized for an abstract, analytical representation that encodes the relationships between object parts, while the RH is specialized for form-specific, holistic, configural processing (e.g., Farah, 2004).

Although the terms “configural” and “holistic” are widely used to describe RH visual processing, little attention has been paid to specifying exactly what they mean. Wouldn’t any high-level visual representation have to be based on more low-level representations? If so, how could there be a “holistic” encoding that isn’t based on constituent parts? I suggest that it is the nature of parts that varies across the hemispheres. In particular, I propose that RH parts are frequency components (i.e., gratings), whereas LH parts are what we normally think of as parts, such as volumetric units.

It is well known that V1 cells encode spatial frequency and orientation for small portions of the visual field. A 2D Fourier transform can be carried out on any image to decompose it into a sum of sinusoidal gratings of different frequencies and orientations. Thus the spatial frequency information encoded in V1 could support two different types of representations; it could be used for edge detection in the spatial domain, or it could be used to assemble gratings for a representation purely in the frequency domain. These two representations would require cells with different types of receptive fields. For the spatial encoding, cells that only respond to 1 or 1 ½ cycles at a given frequency are required, as these would correspond to edges or lines. For the frequency encoding, cells that respond optimally to a large number of cycles are necessary. Indeed, both types of cells have been observed in V1 and V2 (von der Heydt, Peterhans & Dursteler, 1992), while cells in V4 show an even stronger preference for either contours or gratings (Hedge & van Essen, 2007).

RH processing

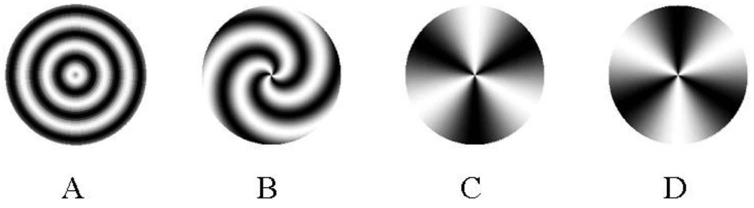

Figure 1: Examples of (A) concentric, (B) spiral, and (C, D) angular gratings. Angular frequency is the same in (C) and (D), but phase is shifted by 180°.

I propose that the RH is specialized for representation in the frequency domain, with grating-sensitive cells preferentially projecting to right inferotemporal cortex. That is, the RH encodes an object as a set of frequency components. This is indeed a holistic representation, as each part (frequency component) spans the entire object. Given that faces are fairly circular, it may be the case that the RH operates on polar, rather than Cartesian, frequencies. That is, the frequency components would be concentric, spiral, and angular gratings, rather than bars. (See Figure 1.)

Indeed, V4 neurons are preferentially responsive to such non-Cartesian gratings (David, Hayden & Gallant, 2006; Hedge & van Essen, 2007). Furthermore, a recent human fRMI study using Gabor fields showed stronger bilateral V4 activation to concentric patterns than to non-concentric patterns, where global and local curvature statistics were controlled (Dumoulin and Hess, 2007). Interestingly, the stereotaxic analysis showed that a ventral RH area anterior to V4 preferred concentric patterns (Dumoulin and Hess, 2007, Fig. 2), as would be expected if the right hemisphere is specialized for processing in the polar-frequency domain, while the homologous LH area did not show a preference for concentric patterns.

Location invariance under a polar-frequency encoding requires the capacity to center the polar coordinate system on the object. This requires cells that respond to polar gratings centered at different retinal locations; such V4 cells have indeed been observed (David et al., 2006). All retinotopic cells representing the same frequency characteristics, but centered at different retinal locations, would then connect to a single abstract representation of those frequency characteristics. Object identification would be carried out over these abstract frequency components, producing location invariance. Multiple objects could be processed seemingly in parallel by multiplexing them over time. That is, the retinotopic (and abstract) frequency components of an object fire synchronously, but asynchronously with those of other objects.

An encoding in the frequency domain entails specifying the amplitude and phase of each frequency component. Amplitude would naturally be represented by activation level, while cells could be tuned to phase in addition to frequency. Note that inverting an object would necessarily change the phase of spiral and angular frequencies with an odd number of cycles. (See Figure 1.) Thus this representation is not rotation-invariant, which would explain the well-known finding that configural processing is impaired for inverted faces.

LH processing

In contrast, I propose that the LH is specialized for representation in the spatial domain, with contour-preferring cells preferentially projecting into left inferotemporal cortex. Thus, LH parts are spatially localized. Again, there would be retinotopic representations of all basic parts centered at different retinal locations. However, a more complex representation is required for local parts than for a holistic encoding. In the holistic case, the relationship between parts (frequency components) is predefined; it is the sum. Thus there is no need to encode part relationships; an object can simply be represented as a set of abstract frequency components. However, this is not sufficient for an analytic encoding, as spatial relationships between parts must be explicitly represented.

I propose that part relationships are represented in a pairwise fashion; an abstract bi-part unit represents a pair of parts in a certain spatial relationship. The location-invariant encoding of an object is then a set of abstract bi-part units. Thus the open-bigram representation in the SERIOL model is taken to be a special case of a general LH representational mechanism. Evidence for this general mechanism comes from single-cell recordings of macaque inferotemporal cortex, where cells were identified that responded preferentially a particular spatial arrangement of a particular pair of features (Brincat & Connor, 2004).

This raises the issue of how such abstract bi-part units are activated. Conventional wisdom would say that bi-part units are activated in parallel. This would require a retinotopic bi-part unit for every possible pair of parts at every possible combination of retinal locations; all retinotopic bi-part units encoding the same spatial relationship between the same pair of parts would then connect to the same abstract bi-part unit. Indeed, implemented parallel models that learn location-invariant object recognition require that every object be presented at a large majority of possible retinal locations (Elliffe, Rolls, & Stringer, 2002), or use the biologically implausible approach of weight sharing, where learning at one location is simply replicated across all locations (Riesenhuber & Poggio, 1999). These measures are necessary to achieve the required repetition of retinotopic part-conjunction detectors.

However, there is an alternative mechanism that doesn't require such repetition. Note that mapping a spatial representation onto a temporal one automatically creates a location-invariant representation; part relationships are encoded in the firing pattern, rather than by retinotopicity. Others have also recognized that a spatial-to-temporal mapping directly achieves location invariance (Buonomano & Merzenich, 1999; Wyss, Konig, & Verschure, 2003). Thus, retinotopic parts may activate a temporal encoding of abstract parts. Pairs of abstract parts would then activate abstract bi-parts, based on firing order. Hence, the manner in which letters and open-bigrams are activated in the SERIOL model is taken to be a special case of a general mechanism for achieving location invariance in an analytical representation.

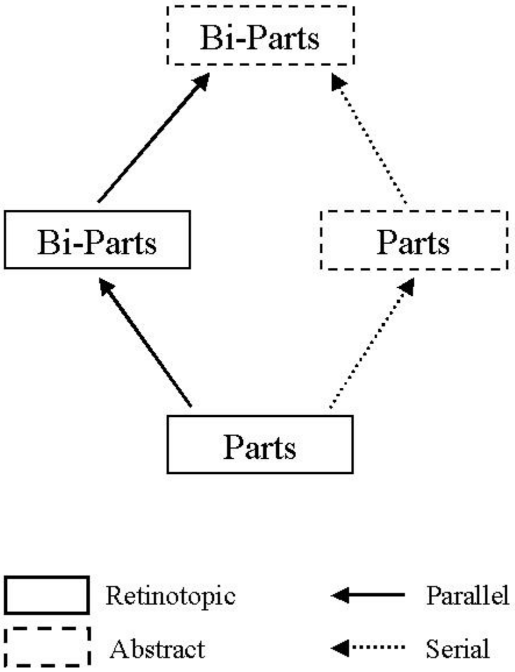

Figure 2: Two possible mechanisms for generating an abstract bi-part representation. On the left, all processing is parallel, which necessitates retinotopic bi-part units. On the right, a temporal encoding of abstract parts replaces the retinotopic bi-parts.

The two mechanisms for achieving an abstract bi-part representation are compared in Figure 2. A temporal encoding allows a location-invariant encoding to emerge at a lower representational level (parts) than does a parallel encoding (bi-parts). Note that it is meaningful to distinguish representations at the level of parts versus bi-parts, as there is a qualitative difference between the two levels. A part represents a spatially contiguous region of an object, while a bi-part encodes pairs of parts that are not necessarily spatially contiguous.

For example, seriality allows an abstract encoding at the level of individual letters in the SERIOL model. In contrast, alternative models of letter-position encoding that propose *parallel* activation of open-bigram units require retinotopic bigram representations (Grainger et al., 2006; Dehaene, Cohen, Sigman & Vinckier, 2005), and do not include an abstract letter level, because there is no way to bind spatial information to abstract letter units under this scheme. However, masked priming experiments indicate that the same letter representation is activated by a given letter at different retinal locations and string positions, implying the existence of abstract letter units (Peressotti & Grainger, 1995). These results are consistent with abstract letter representations in the SERIOL model, and suggest that location invariance in general arises at the level of individual parts.

A temporal encoding requires that the order in which parts are "read out" be invariant. Such invariance could be achieved relative to center of attention (which usually, but not always, corresponds to fixation). That is, the activation gradient caused by attention would cause retinotopic parts nearer the focus of attention to fire before those farther from the focus (Connor, Gallant, Preddie & van Essen, 1996). Stored object representations in long-term memory could record multiple read-outs originating from different canonical points, such as from the center and top of the object.

This requirement for the temporal encoding to originate at the focus of attention implies that it should only be possible to construct an analytical encoding of an attended object. This is consistent with the results of a priming study of non-symmetric images of objects. An attended prime facilitated subsequent processing of its left-right reflection, while unattended prime did not prime its reflection (Stankiewicz, Hummel & Cooper, 1998). An analytic read-out from the center of an object would create a reflection-invariant representation, while a holistic representation based on angular / spiral phase would not. Hence, this finding is consistent with the assumption that attention is necessary to form an analytical representation; in the absence of attention, only a holistic representation is formed. This account predicts that attended, but not ignored, images should generalize over a distortion that destroys configural relationships but leaves structural relationships largely intact, such as splitting an image in half. Indeed, attended split images primed their intact counterparts but unattended split images did not (Thoma, Hummel & Davidoff, 2004).

In summary, I propose that because an analytical representation requires encoding relationships between parts, spatial relationships are mapped onto a temporal encoding to efficiently induce location invariance. This mapping requires an attention gradient, and only allows analysis of one object at a time. In the RH, the parts of an object can fire synchronously because the relationship between parts is predefined. Different objects fire asynchronously to provide seemingly parallel processing of multiple objects, but firing relationships between objects do not encode information. (Rather, spatial relationships between objects are encoded by the parietal lobes.)

Thus I propose that time is used for different purposes in the two hemispheres. In the LH, time encodes spatial relationships within an object. In the RH, time is used to separate the representations of multiple objects, as within-object spatial relationships do not have to be represented due to the nature of the frequency encoding. As discussed below, I suggest that this difference is crucial for understanding the nature of pure alexia.

Extended SERIOL Model

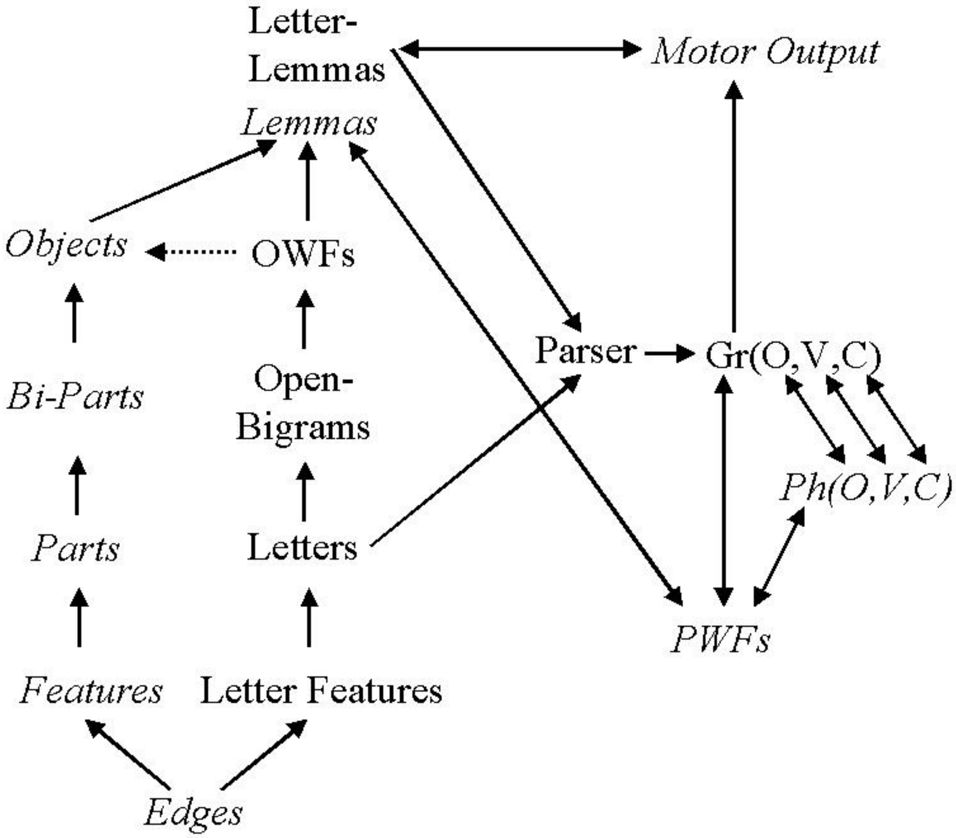

Figure 3: Extended SERIOL model and its relationship to innate LH visual and phonological processing. Innate representations are shown in *Italics*, while representations that are specific to letter strings are shown in **Normal Face**. The visual/ventral route is shown on the left, and the phonological/dorsal route on the right. Abbreviations: OWFs – Orthographic Word Forms; PWFs – Phonological Word Forms; Gr – Graphosyllables; Ph – Phonological syllables; (O,V,C) – (Onset, Vowel, Coda).

The original SERIOL model describes a specialization of LH visual analytical processing for letter strings, as illustrated in the left half of Figure 3. Next the model is extended to include the dorsal route, shown in the right half of Figure 3. The following is a refinement of some of the ideas first presented by Whitney and

Cornelissen (2005). The dorsal route is not yet specified in as much neural detail as the ventral route, but the following characterization suffices for the present purposes.

The word units of the original model correspond to orthographic word forms (OWFs), which are analogous to a visual, parts-based representation of an object. In the extended model, the OWFs then connect to lemmas in temporal cortex, which encode lexicosemantic information across modalities. Bi-directional connections link lemmas and phonological word forms (PWFs) in the temporoparietal area.

The edge, feature, and letter layers of the original model are common to both routes. Note that the letter layer is a mid-level visual representation. In addition, there would be a more conceptual letter representation in the lemma layer, called *letter-lemmas*. Letter-lemmas allow conscious access to letter information. On the input side, such high-level letter representations can be activated by individual letters or oral letter names; on the output side, letter-lemmas are activated by writing, typing, or oral spelling.

On the dorsal route, a sequence of letters is parsed into a graphosyllabic representation, which encodes graphemic onsets, vowels, and codas. It is assumed that the parser can take input in the form of the serial encoding from the letter level, or as a sequence of letter-lemmas. The slower processing required for the latter type of input could be learned via the feedback loop from letter-lemma output. Graphosyllables are mapped to syllables, which encode phonological onsets, vowels, and codas; syllables connect bidirectionally to PWFs. The mapping between graphosyllables and syllables supports sublexical phonological assembly.

I propose that there are also direct bi-directional connections between graphosyllables and PWFs. Thus, the dorsal route encodes lexical information, which can be used for both input and output. Spelling is carried out by connections from PWFs to graphosyllables to serial activation of letter-lemmas for motor output. Orally-spelled words are recognized via connections from letter-lemmas to graphosyllables to PWFs. Visually-presented strings activate the pathway from letters to graphosyllables to PWFs.

Therefore, I propose that visual word recognition employs two different high-level orthographic representations. On the ventral route, open-bigrams provide a flat relative-position encoding. On the dorsal route, graphosyllables provide a hierarchical slot encoding. This implies that the dorsal route actually constructs a more veridical orthographic representation than the ventral route, as open-bigrams introduce ambiguity while graphosyllables precisely encode orthographic structure. For example, the string BLET shares many open-bigrams with BELT (i.e., ***B, BE, BL, BT, LT, ET, and T***, where * represents a boundary), while BLET mismatches BELT in orthographic onset and coda (i.e., **BL** vs **B**, and **T** vs **LT**). Thus BELT and BLET are more different under a graphosyllabic encoding than an open-bigram representation.

Therefore, the dorsal route should be better able to detect misordered letters than the ventral route. This is consistent with the results of lexical-decision experiments (Frankish & Turner, 2007) using nonwords formed by transposing adjacent letters of a real word (TL nonwords). Pronounceable TL nonwords (such as SONUD) were *easier* to reject than unpronounceable ones (such as SIHGT). Note that this pattern radically differs from non-TL nonwords, where pronounceable items are *more difficult* to reject than unpronounceable ones, due to increased word-likeness. The TL results indicate that the dorsal route produced a veridical parse and phonological encoding of a pronounceable TL nonword, which mismatched and inhibited lexical activation of the base word from the less precise ventral route (Frankish & Turner, 2007). In the case of an unpronounceable TL nonword, the dorsal route cannot parse the stimulus, so there is no mismatching phonological representation to inhibit the ventral route, leading to increased false positives.

The proposal that the dorsal route contributes lexical information to visual word recognition is consistent with the fact that damage near the left occipito-temporo-parietal junction causes reading errors for words (as well as pseudowords). Lesions to this area cause visual and semantic errors, with positive effects of imageability and frequency (Crisp & Lambon Ralph, 2006; Friedmann, 1996; Jefferies, Sage & Ralph, 2007). In this case, lexical activation would be driven by the less precise open-bigram representation, explaining visual errors (Crutch & Warrington, 2007). This analysis suggests that a particular deficit in encoding letter order should be observed. Indeed, patients with left occipitoparietal lesions display letter-position dyslexia (Shalev, Mevorach, & Humphreys, 2008; Friedmann & Gvion, 2001). Without constraints from the direct activation of the correct PWF via the dorsal route, spreading activation within the semantic network (i.e., between lemmas) could cause related higher-frequency lemmas to become more activated than the correct lemma, yielding semantic errors, particularly for low frequency stimuli. Given that OWFs and visual object forms are taken to occupy the same brain region (i.e., left anterior fusiform gyrus), there may be direct connections between OWFs and their corresponding object forms (as shown by the dotted arrow in Fig. 3), which would account for the influence of imageability.

In contrast, rare cases of acquired pure phonological dyslexia have been observed, where word reading is normal but pseudoword reading is not (Caccappolo-van Vliet, Miozzo & Stern, 2004). This shows that damage to the sublexical system doesn't necessarily cause errors for words, supporting the idea that the above semantic and visual errors are due to the absence of lexical information from the dorsal route, rather than the presence of noisy sublexical information. Pure phonological dyslexia would arise when connections between graphosyllables and syllables are damaged, but connections between graphosyllables and PWFs remain largely intact.

If the dorsal route can handle both lexical and sub-lexical processing, of what use then is the open-bigram encoding on the ventral route? It has the advantage of speed. Visual information zooms along the fusiform gyrus, allowing rapid lexical access. In contrast, the dorsal route involves connections across longer distances, as well as the more complex computation of syllabic structure. Thus the ventral route allows fast access to easily activated lexical items (i.e., high frequency words), while the slower dorsal route contributes to recognition of lower-frequency and novel words via more precise lexical and sublexical analysis.

In summary, the key proposals are the following. The dorsal route constructs its own high-level orthographic encoding, based on graphosyllables. Bi-directional connections between graphosyllables and PWFs encode lexical information. Graphosyllables are activated by the graphosyllabic parser, which takes serial letter input. This input may be very rapid (via visual string processing) or slow (via an oral sequence of letter names).

Pure Alexia

It is now straightforward to explain why letter-by-letter readers do so. Left occipitotemporal damage would wipe out the ventral route. However, it would spare connections between graphosyllables, PWFs, and lemmas; lexical access could be carried out along this pathway. A serial letter encoding is required to activate the spared graphosyllabic representations. However, left occipitotemporal damage would destroy the capability to produce a very rapid serial encoding of visually presented strings, as this type of representation is specific to the LH.

If the RH were to process the string as a single object, it would produce a holistic representation that does not encode letter identities. Such a representation is clearly not suitable for activating the graphosyllabic system. Therefore the RH must process the letters as individual objects. This would produce asynchronous firing of RH letter representations, because objects are multiplexed over time. However, this multiplexing would not yield serial left-to-right activation, as the RH firing pattern does not encode spatial information. Therefore the only way to provide the required serial input to the spared graphosyllabic system is via overt letter-by-letter processing.

Discussion

In the model, a serial letter representation provides input to the ventral and dorsal routes. It is increasingly becoming accepted that phonological assembly normally proceeds sequentially on the dorsal route (e.g., Carrieras, Ferrand, Grainger & Perea, 2005; Coltheart et al., 2001; Perry, Ziegler & Zorzi, 2007), in line with the proposed serial letter encoding. However, it is generally assumed that the ventral route operates in parallel, primarily due to the lack of a length effect for words in lexical decision and naming (Frederiksen & Kroll, 1976; Weekes, 1997). Because the present account of pure alexia crucially depends on the proposal that serialization normally occurs in left occipitotemporal cortex (and hence includes the ventral route), it is important to establish that the absence of a lexical length effect does not preclude a serial encoding.

As pointed out by Whitney and Lavidor (2004), serial processing could fail to yield a length effect if increased length also has a counterbalancing facilitative effect. Additional letters may reduce the time required for the lexical network to settle after the final letter fires (as compared to shorter words). Thus increased letter-processing time and decreased lexical-settling time may cancel each other out, giving no length effect despite a serial encoding.

In fact, a recent EEG study of length effects in lexical decision provides direct support for this scenario (Hauk & Pulvermuller, 2004). Word length had no effect on reaction times, but yielded complementary effects on EEG amplitude at different time periods. From 100 to 125 ms post-stimulus, longer words gave increased activity;

from 150 to 360 ms, longer words yielded *decreased* activity. These results are entirely consistent with the proposal that longer words induce more efficient processing at the lexical stage, canceling out increased processing time at the letter stage.

A subsequent ERP experiment revealed more detail about the early increased activity for longer words (Hauk, Davis, Ford, Pulvermuller & Marslen-Wilson, 2006). In this study, the increased activity extended from 90 ms to 220 ms, while the decreased activity began after 300 ms. Crucially, the length effect was initially right lateralized (at 90 ms) and then became left lateralized (at 200 ms). This shift demonstrates that the early length effect is not simply due to increased visual angle for longer words, because such an effect should be symmetric. Moreover, it supports the claim of serial processing. The first half of a centrally-fixated word falls in the left visual field (LVF) and is projected to the right hemisphere (RH), while second half falls in the RVF and is projected to the LH (Hunter, Brysbaert & Knecht, 2007). Hence, the shift of lateralization from right to left is consistent with the letters of the first half of the word being processed prior to those of the second half.

For lexical decision under unilateral presentation, there is a length effect for words in the LVF, but not the RVF (Young & Ellis, 1985). The LVF length effect and the phenomenon of LBL reading have often been interpreted as indicating that the RH can only support serial lexical access, while the LH provides parallel lexical access. However, as discussed by Barber & Kutas (2007), this is inconsistent with the brain-imaging evidence for *prelexical* left lateralization, which indicates that there is a single mechanism of lexical access within the LH.

Hence, any LVF and RVF differences must occur prior to this confluence in the LH. That is, VF differences must arise at a visual-orthographic level, and a single method of lexical access must produce both the presence and absence of a length effect. This easily explained by the serial account. If the facilitative and inhibitory influences exactly cancel each other out, there is no length effect. If the balance is altered such that the inhibitory influence outweighs the facilitative one, a length effect emerges. This could occur via a decreased facilitative and/or increased inhibitory effect, due to degradation of the visual / orthographic encoding. Indeed, a length effect emerges under RVF presentation when the early visual encoding is compromised via left occipital TMS (Skarratt & Lavidor, 2006).

For LVF presentation, the SERIOL model implies that the normal length effect is due to a non-optimal orthographic activation pattern (locational gradient), which reduces the facilitative effect of additional letters and reveals the underlying seriality. This non-optimality arises from the conversion of the increasing acuity gradient into the decreasing locational gradient. Based on this analysis, we have succeeded in abolishing the LVF length effect via positional adjustments of contrast level specifically designed to correct the LVF locational gradient (Whitney & Lavidor, 2004). Such a visual manipulation could not have converted serial processing into parallel processing. Hence, the absence of a length effect cannot be used to infer the absence of serial processing.

If the facilitative influence were to outweigh the inhibitory one, there could be a reverse (facilitative) length effect, even under serial processing. This explains the surprising results of a recent analysis of behavioral length effects based on the English Lexicon Project, an on-line database of lexical-decision reaction times for over 40,000 (centrally-presented) words (Balota, Cortese, Sergent-Marshall & Spieler, 2004; New, Ferrand, Pallier & Brysbaert, 2006). Once the effects of frequency, number of syllables, and orthographic neighborhood size were factored out, reaction times *decreased* with increasing length for words of three to five letters (and remained constant for words of five to eight letters). Thus there is actually a facilitative length effect for shorter words. (This effect had not been observed in previous studies because neighborhood size had not been controlled.) It appears that the facilitative influence outweighs the cost of processing more letters at shorter word lengths.

Indeed, simulations of the open-bigram and word layers of the SERIOL model have successfully replicated this pattern. Despite *serial* activation of open-bigrams, settling time at the lexical level *decreased* with increasing length for words of three to five letters, and remained constant for words of five to seven letters (Whitney, in press). For shorter words, the increased amount of information from additional letters allowed the lexical network to settle faster, outweighing the increased processing time at the open-bigram level.

In sum, reaction times in lexical processing are not solely a function of long it takes to activate letter representations. These simulations and experimental results indicate that a serial encoding of letter order is not inconsistent with an absence of an inhibitory length effect.

Moreover, only a serial representation can explain an interesting pattern of letter perceptibility results, as discussed in more detail by Whitney (in press). For very brief presentations (< 100 ms) of four- or five-letter strings, there is a strong initial-letter advantage, but the final letter is the least well perceived (e.g., Humphreys, Evett, & Quinlan, 1990; Tydgat & Grainger, 2007). At longer durations (> 100 ms), the well-known final-letter advantage emerges – the last letter is perceived better than the internal letters (e.g., Mason, 1982). This is easily explained under a serial encoding. For very brief presentations, there is insufficient time to reliably reach the final letter. When exposure duration is increased, the final letter is activated, and can then fire more robustly than the internal letters because it is not inhibited by a subsequent letter. This produces a final-letter advantage. It is unclear how these patterns could be explained under purely parallel processing.

In conclusion, I propose that LBL reading emerges because letter order is normally represented serially and a serial encoding is required for lexical access; this very rapid serial encoding is normally induced in the left fusiform gyrus and cannot be produced by the RH. Thus I propose the exact opposite of the standard view of pure alexia. The problem is not that the RH is incapable of processing letters in parallel; the problem is that the RH cannot automatically process letters serially. Unfortunately, if this analysis is correct, it implies that little can be done to ameliorate letter-by-letter reading.

References

Baker, C.I., Liu, J., Wald, L.L., Kwong, K.K., Benner, T. &, Kanwisher, N. (2007) Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, 104, 9087-9092.

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

Barber, H.A. & Kutas, M. (2007) Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Research Reviews*, 53, 98-123.

Beeson, P.M., Magloire, J.G. & Randall, R.R. (2005) Letter-by-letter reading: Natural recovery and response to treatment. *Behavioral Neurology*, 16, 191-202.

Behrmann, M, Shomstein, S.S., Black, S.E. & Barton, J.J. (2001) The eye movements of pure alexic patients during reading and nonreading tasks. *Neuropsychologia*, 39, 983-1002.

Brincat, S. L. and Connor, C.E. (2004) Underlying principles of visual shape selectivity in posterior inferotemporal cortex, *Nature Neuroscience*, **7**, 880–886.

Buonomano, D.V. & Merzenich, M. (1999) A neural network model of temporal code generations and position-invariant pattern recognition. *Neural Computation*, 11, 103-116.

Caccappolo-van Vliet, E., Miozzo, M. & Stern, Y (2004) Phonological dyslexia: A test case for reading models. *Psychological Science*, 15, 583-590.

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

Coltheart, M. Rastle, K. Perry, C. Langdon, R. & Ziegler, J. (2001) DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204-256.

Cohen, L., Henry, C., Dehaene, S., Molko, N., Lehéricy, S., Martinaud, O., Lemer, C., and Ferrieux, S. (2004). The pathophysiology of letter-by-letter reading. *Neuropsychologia*, 42, 1768-1780.

Cohen, L., Martinaud, O., Lerner, C., Lehericy, S., Samson, Y., Obadia, M., & Slachevsky, A. (2003). Visual word recognition in the left and right hemispheres : Anatomical and functional correlates of peripheral alexias. *Cerebral Cortex*, 13, 1313-1333.

Connor, C.E., Gallant, J.L., Preddie, D.C. & van Essen (1996) Responses in area V4 depend on the spatial relationship between stimulus and attention. *Journal of Neurophysiology,* 75, 1306-1308.

Crisp, J. & Lambon Ralph, M.A. (2006) Unlocking the nature of the phonological-deep dyslexia continuum: the keys to reading aloud are in phonology and semantics. *Journal of Cognitive Neuroscience,* 18, 348-362.

Crutch, S.J. & Warrington, E.K. (2007) Word form access dyslexia: understanding the basis of visual reading errors. *Quarterly Journal of Experimental Psychology*, 60, 57-78.

David, S.V. Hayden, B.Y. & Gallant, J.L. (2006) Spectral receptive field properties explain shape selectivity in area V4. *Journal of Neurophysiology*, 96, 3492-3505.

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

Farah, Martha, J. (2004) Visual Agnosia. MIT Press: Cambridge, MA.

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

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

Friedman, R.B. (1996) Recovery from deep alexia to phonological alexia: Points on a continuum, *Brain and Language,* 52, 114–128.

Friedmann, N. & Gvion, A. (2001) Letter position dyslexia. *Cognitive Neuropsychology*, 18, 673-696.

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

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

Hauk, O, Davis, M.H., Ford, F. Pulvermuller, F. & Marslen-Wilson, W.D. (2006) The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage*, 30, 1383-1400.

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

Hedge, J. & van Essen, D.C. (2007) A comparitive study of shape representation in macaque visual areas V2 and V4. *Cerebral Cortex*, 17, 1100-1116.

Humphreys, G. W., Evett, L. J. & Quinlan, P. T. (1990). Orthographic processing in visual word identification. *Cognitive Psychology*, 22, 517- 560.

Jeffries, E., Sage, K. & Ralph, M.A. (2007) Do deep dyslexia, dysphasia, dysgraphia share a common phonological impairment? *Neuropsychologia*, 45, 1553-1570.

Johnson, R.L. & Rayner, K. Top-down and bottom-up effects in pure alexia: evidence from eye movements. *Neuropsychologia*, 45, 2246-2257.

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

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

McCandliss, B.D., Cohen, L. & Dehaene, S (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Science*, 7, 293-299.

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

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

Perry, C., Ziegler, J.C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, 114, 273-315.

Riesenhuber, M. & Poggio, T. (1999) Hierarchical models of object recognition in cortex. *Nature Neuroscience,* 2, 1019–1025.

Shalev, L., Mevorach, C. & Humphreys, G.W. (2008) Letter position encoding in attentional dyslexia. *Neuropsychologia*, in press.

Skarrat, P.A. & Lavidor, M (2006) Magnetic stimulation of left visual cortex impairs expert word recognition. *Journal of Cognitive Neuroscience*, 18, 1749-1758.

Stankiewicz, B., Hummel, J.E. & Cooper, E.E. (1998) The role of attention in priming for left-right reflections of object images; evidence for a dual representation of object shape. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 732-744.

Thoma, V., Hummel, J.E. & Davidoff, J. (2004) Evidence for holistic representations of ignored images and analytic representations of attended images. *Journal of Experimental Psychology: Human Perception and Performance,* 30, 257-267.

Tydgat, I. & Grainger, J. (2007). Serial position effects in the identification of letters, digits, and symbols. In Proceedings of XVth Meeting of the European Society for Cognitive Psychology (p. 123).

von der Heydt, R., Peterhans, E. & Dursteler, M.R. (1992) Periodic-pattern-selective cells in monkey visual cortex. *Journal of Neuroscience*, 12, 1416-1434.

Weekes (1997) Differential effect of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology*, 50, 439-456.

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

Whitney, C. (2004). *Investigations into the Neural Basis of Structured Representations*. Doctoral Dissertation. University of Maryland.

Whitney, C. (in press). Supporting the serial in the SERIOL model. *Language and Cognitive Processes.*

Whitney, C. & Cornelissen, P. (2005). Letter-position encoding and dyslexia. *Journal of Research in Reading*, 28, 274-301.

Whitney, C. & Lavidor, M. (2004). Why word length only matters in the left visual field. *Neuropsychologia*, 42, 1680-1688.

Wyss, R., Verschure, P.F., & Konig, P. (2003) Invariant representations of visual patterns in a temporal population code. *Proceedings of the National Academy of Sciences*, 100, 324-329.

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