1420

British Journal of Psychology (2006), 97, 191–216

The British Psychological Society

www.bpsjournals.co.uk

The orthographic uniqueness point and eye movements during reading

Brett Miller[1]*, Barbara J. Juhasz[2] and Keith Rayner[2]
[1]Haskins Laboratories, USA
[2]University of Massachusetts, USA

Recent research found that naming and lexical decision times for words with an early orthographic uniqueness point (OUP) were faster than for words with a late OUP (Kwantes & Mewhort, 1999a; Lindell, Nicholls, & Castles, 2003). A word's OUP corresponds to the letter position in the word where that word is differentiated from other words. These results have been presented as evidence for sequential letter processing in visual word recognition (Kwantes & Mewhort, 1999a). In two experiments, we attempted to extend these results to a more natural reading situation by recording participants' eye movements. Readers read sentences with early or late OUP words embedded in them. In both experiments, we manipulated the amount of parafoveal information available during reading. Readers did not show any consistent benefit for reading words with an early OUP regardless of the amount of preview available. Our results are at odds with the naming and lexical decision data and prove problematic for models that predict OUP effects.

When developing models of visual word recognition, it is necessary to posit the nature of the mechanisms involved in letter identification and how this letter information is used by the model. Typically, models primarily fall into two camps: *hybrid-type* models that have some serial and parallel components (e.g. Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Kwantes & Mewhort, 1999b) and *fully parallel* models (e.g. Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McCleland, 1989).

In the present context, *parallel letter processing* refers to the processing of multiple letters in a word at the same time while *serial letter processing* refers to the processing of one letter at a time. The more general distinction between serial and parallel processing could occur at a variety of different processing steps that are involved in processing letter information in order to identify a word. For instance, one could imagine a hybrid letter processing account that assumes parallel feature and letter activation but a serial readout of the letter identities used to identify the target word

*Correspondence should be addressed to Brett Miller, Haskins Laboratories, 300 George St. New Haven, CT 06511, USA (e-mail: miller@haskins.yale.edu)

DOI:10.1348/000712605X66845

(e.g. Kwantes & Mewhort, 1999b). To limit our discussion somewhat, we will focus on accounts of visual word recognition that posit predictions regarding the role of a word's orthographic uniqueness point, which is the point in a word where the preceding letters are sufficient to uniquely identify the word.

Recently, Kwantes and Mewhort (1999a) reported data suggesting that when people retrieve a word from memory, they search their lexicon serially and sequentially using the letters in a target word from left to right when naming words in English. Participants named words that varied in their orthographic uniqueness point (OUP). To illustrate, consider the early OUP word *actress*. The first four letters, '*actr*___', are sufficient to identify the word because *actr* can only form the word *actress* or a derivative of it. In contrast for the late OUP word *cartoon*, participants must identify all letters up through the last '*o*' before differentiating *cartoon* from all lexical competitors. Kwantes and Mewhort found that participants named early OUP words 29 ms faster than late OUP words (Experiment 1). They argued that faster naming times for early OUP words provided strong evidence that participants process letters in a serial, sequential nature when naming words.

Kwantes and Mewhort (1999b) posited that readers first need to identify the letters and then retrieve the target word from memory. Their model (LEX) is based on the MINERVA model (Hintzman, 1984) and assumes that a reader searches the lexicon for a given word using the word's letter identities. To simplify the search process, the reader searches memory sequentially letter-by-letter until the word can be identified uniquely from its competitors. It is this serial probing that generates an OUP effect because fewer letters and hence fewer searches would need to occur on average for words whose spelling pattern is unique at the beginning portion rather than later in the word (see also Kwantes & Mewhort, 1999a for short review). In the context of this model, seriality occurs not at the letter identification stage, but rather when readers attempt to match the letter information to known words.

In contrast, fully parallel models of visual word recognition should not produce an OUP effect. To ground this argument, take the model by Plaut *et al.* (1996). In this account, letters are identified in parallel as featural information becomes available. A letter candidate receives greater activation as additional, consistent information is processed and activation becomes reduced for other letter candidates. This process occurs across all letter positions in parallel. Each letter candidate activates candidate words in parallel that possess the positionally consistent letter. Over time, activation increases for the target word until it reaches criterion for identification. This type of an account should not predict a difference for a word that possesses an early OUP versus a late OUP. Without an extension of the model, there is no mechanism included in this model that gives preference to the processing of the initial letter positions over other letter positions making it impossible to generate an OUP effect unless the materials contained a confound that the model might be sensitive to.

OUP effects provide some primary evidence for serial letter processing when naming words. Other evidence for serial processing during word recognition comes from Roberts, Rastle, Coltheart, and Besner (2003) who reported data showing that the point of irregularity affected readers' naming times for irregular words. Roberts *et al.* interpreted their findings within the context of the DRC model (Coltheart *et al.*, 1993, 2001), which includes a parallel, lexical route and a serial, non-lexical route. In this account, if the irregularity occurs early in a word, then competition occurs between the two routes because the outputs generated early on would be inconsistent. However, for irregularities occurring later in a word, the lexical route will be more likely to finish processing first making the inconsistent output from the nonlexical route irrelevant.

One might also expect to find serial processing in reading for reasons relating to perceptual span and attention (Henderson & Ferreira, 1990; Rayner, 1998). Visual acuity constraints limit the information that a reader is able to process during a fixation, which could prevent readers from processing all the letters in a word in parallel. In fact, the word identification region, a subsection of a reader's perceptual span, covers only about eight character spaces to the right of fixation (McConkie & Zola, 1987; Underwood & McConkie, 1985) making it unlikely that readers process letters in words in a strictly parallel fashion; this is particularly true for longer words (for a similar argument, see Bertram & Hyönä, 2003; Hyönä & Pollatsek, 1998; Plaut *et al.*, 1996). This argument should be clear for longer words where, regardless of the location of a single fixation on a word, it would be unlikely that all the letters would fit in the word identification region of the perceptual span. This does not imply that people are not processing information about the remaining letters, but rather only that it is unlikely that information sufficient to identify all letters would be available during a single fixation for these words. This limitation does not exclude the possibility that letter identification occurs in parallel within the word identification region. Finally, at the interword and intraword level, a prominent model of eye movement control in reading, the E-Z reader model, posits that attentional shifts occur serially in reading (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2003). Thus, the presence of some degree of seriality during reading seems viable.

We had three primary motivations for conducting Experiment 1. First, we wanted to examine whether we could find evidence for serial letter processing in reading by extending the OUP results to a more natural reading environment by recording participants' eye movements during sentence reading. Secondly, we wanted to verify that OUP effects extend to a broader range of words; hence, we used stimuli derived from a different corpus than those used in previously reported experiments (Kwantes & Mewhort, 1999a; Lindell *et al.*, 2003; Pacht, 2003). Finally, we wanted to examine the role of parafoveal processing with respect to the OUP. Specifically, we manipulated the amount of parafoveal information that was available from a target word prior to its fixation and also varied whether the information available was or was not consistent with the OUP for those target words.

Experiment 1 compared reading times for early OUP words to those for late OUP words. We incorporated a parafoveal preview manipulation that allowed precise control of the preview available to the readers before they fixated the target word (see Fig. 1; for discussion, see Rayner, 1975, 1998). Three preview conditions were included. Readers received a preview that was either (1) identical to target, (2) consistent with the target through the first four letters only or (3) a preview that had no overlapping letters with the preview. If readers process letters in a serial, sequential fashion, then they should read the early OUP words faster than the late OUP words regardless of preview condition. If readers process letters in a parallel fashion, then the predictions differ (see below).

To illustrate how these predictions vary, we turn to an extension of a model proposed by Plaut *et al.* (1996). For longer words, they suggested that parallel processing might be limited to an attentional window covering only part of a word. Extending this argument to eyetracking, the saccade launch site could affect whether OUP effects are found. The logic is that for cases where the reader makes a saccade from a location close to the target, the target word's initial letters would probably be inside of

Early OUP – Preview-Denied

*

1a. Jeanette saw a baby thqjzwp for the first time at the zoo.

*

1b. Jeanette saw a baby giraffe for the first time at the zoo.

Early OUP – Partial Preview

*

2a. Jeanette saw a baby girazwp for the first time at the zoo.

*

2b. Jeanette saw a baby giraffe for the first time at the zoo.

Early OUP – Full Preview

*

3a. Jeanette saw a baby giraffe for the first time at the zoo.

*

3b. Jeanette saw a baby giraffe for the first time at the zoo.

Figure 1. Examples of preview information available to a reader contingent upon eye position. This figure illustrates the information available parafoveally to the reader before they fixate or skip over the target word ('a' sentences above) and also foveally after they fixate or skip the target word ('b' sentences above). The * corresponds to the reader's fixation location. The examples are broken into three groups by preview condition type: 1. Preview denied condition, 2. Partial Preview condition and 3. Full Preview condition.

this attentional window, which might correspond roughly to the word recognition region of a reader's perceptual span. This seems reasonable if we assume that high fidelity letter information is only available within this attentional window. With this in mind, the type of preview information available would dictate whether an OUP effect might occur. For the preview denied condition, readers would not be expected to read the early OUP words faster than the late OUP words because no useful parafoveal preview was available. In the partial and full preview cases, the reader might obtain substantial parafoveal preview of the initial letters. This parafoveal processing could generate OUP-like effects because the initial letters would be identified before later letters truncating the list of lexical candidates further for the early OUP.

For cases where readers launch their saccade into the target word from locations far away from the target word, no effect for OUP is expected according to this account. By far launch sites, we are referring to saccade locations where it is unlikely that all the target words' initial letters would be within the word identification window. For the far launch sites, we chose launch sites that were five or more characters away from the target region. The motivation for this cut-off derives from the size of the word identification region. Since the word identification region is about eight characters to the right of fixation, the choice of five characters or more from the target region should insure that the OUP is outside this region and hence receiving minimal parafoveal processing. Without substantial parafoveal preview of the target, the initial letters would not receive much, if any, of a preprocessing boost before fixating the target. Upon fixating the target, the reader would process the target letters in parallel and hence not be sensitive to a word's OUP.

Our experiment is also suited to test the split processing model (Shillcock, Ellison, & Monaghan, 2000). Essentially, the model partitions a word's input into two parts roughly at the informational midsection and one part goes to each hemisphere of the brain. The word parts may or may not correspond to a word half; this is particularly unlikely for very long words. Each word part generates a separate set of lexical competitors. Over time, these candidate sets are reduced until the reader identifies the word presented. This model leaves open a number of questions making it difficult to derive precise predictions.

First, this account does not specify when the two hemispheres will begin to share information that could be used to reduce the lexical competitor sets. One could imagine a number of scenarios. On one extreme, one might claim that the two hemispheres do not communicate until each candidate set is fully resolved. In this case, if the reader is unable to resolve a candidate set fully with the initial information presented to either hemisphere the reader would need to refixate the word so that the information necessary to resolve each candidate set would be available. This type of processing seems unlikely, particularly given the efficiencies that could be gained by having some degree of cross-talk before a lexical candidate set is fully resolved.

Second, one could imagine a case where the two hemispheres communicate immediately upon receiving input from the visual system. Shillcock *et al.* (2000) specifically discount this scenario. The logic of their argument is that if there is immediate sharing of information about the inputted word before the reader begins higher level computations, then one would not expect to find differential processing abilities when presenting information to a single hemisphere, but in fact these differential propensities exist (see Shillcock *et al.*, 2000, for details).

Finally, this leaves us with an account that involves some degree of inter-hemispheric cross-talk as each hemisphere processes its visual input concurrently. This is the assumption that we use when deriving the model predictions. We assume that the hemispheres share information as the candidate sets are reduced. If this is the case, then Shillcock *et al.*'s account should predict a benefit for words possessing an early OUP because the words should have fewer competitors for the word initial portion than late OUP words. For the early OUP words, the candidate set for the right hemifield (initial portion of the word) will converge on the target word, but this will not be the case for the late OUP words and this should result in faster reading times assuming that word length encoding is not precise (Fischer, 2000).

This prediction is based on two assumptions: first, on average, there is no difference in the informational distribution for the ending portion of the target word between the early and late OUP words. The reason this is a potential concern is that if the late OUP words were informationally denser at the end portion of the word, then readers might converge on a solution (the target word) for this second part faster than in the early OUP case. This could reduce any processing difference that might be present between the early and late OUP words. We have no reason to believe that, on average, the ending of our early and late OUP words will differ in terms of their informativeness. Additionally, there is evidence for a processing preference for the beginning of a word (e.g. Brysbaert, 1994; Pynte, Kennedy, & Murray, 1991; Rayner, McConkie, & Zola, 1980). If this is the case, then one might expect faster processing for the early OUP words offsetting any potential discrepancies in the informativeness of the word ending.

Second, we assume that the informational split is approximately optimal for our words so that the initial letters of the word will be processed by the right hemifield. We have reason to believe this to be the case. Shillcock *et al.* (2000) presented simulations showing that their model generally converges on a fixation point at the centre or slightly to the left of centre for words ranging from 6 to 10 characters. If one looks at the average initial landing position for Experiments 1 and 2, it is clear that our data are consistent on average with this positioning.

In terms of our experimental conditions, the preview denied condition provides the clearest predictions for the split processing model because there is no useful letter identity information available to the reader in the parafovea, which could affect the model's parsing predictions. Also, we emphasize the single fixation duration data when interpreting our findings with respect to this model. In these cases, the reader does not refixate the target on first pass. This should insure that the initial parsing of the target word into two parts only occurs once, which may not be the case for measures that allow for first pass refixations on the target word.

EXPERIMENT 1: OUP EFFECTS IN SENTENCE READING

Method

Participants

Participants in the experiment were 42 University of Massachusetts community members. They either received course credit or were paid $8 per hour. All participants were native speakers of American English and had normal vision or wore contact lenses.

Apparatus

A Fourward Technologies Dual Purkinje eye tracker (Generation V), which has a resolution of less than 10 minutes of arc was used to record participants' eye movements. The sentences were displayed on a 15-inch NEC MutliSync 4FGe monitor. Eye movements were measured from the right eye but viewing was binocular. Participants were seated 61 cm from the computer where one degree of visual angle equalled approximately 3.8 characters.

Procedure

Before the experiment, a bite bar was prepared for participants and they received instructions detailing the procedure. A single-line calibration routine was performed and its accuracy was checked after each sentence. When inaccurate, a new calibration was performed. Comprehension was checked on approximately 15% of the trials during the experiment. On average, accuracy was over 94%. Before the experimental trials began, participants read six practice sentences to acclimate to the task.

Stimuli

Fifty-four pairs of early and late OUP words were chosen so that they were matched on several variables (see Table 1). Early and late OUP words were selected from the MRC psycholinguistic database (Coltheart, 1981). The early OUP words had an average OUP

Table 1. Stimuli properties as a function of OUP condition (early versus late) in Experiment 1 including average OUP, word frequency (freq.), initial trigram frequency (trigram freq.), word length (length), number of syllables (syll.), familiarity (fam.), target word goodness-of-fit (GOF) in the sentence context and predictability from prior sentence context (predict).

	OUP	Freq.	Trigram freq.	Length	Syll.	Fam.	GOF	Predict
Early	4.0 (0)	7.7 (9.5)	24.8 (39.8)	7.4 (1.1)	2.5 (0.64)	5.8 (0.75)	5.6 (0.89)	0.37% (1.91)
Late	6.8 (0.7)	8.0 (8.3)	26.1 (30.8)	7.6 (1.0)	2.5 (0.54)	5.7 (0.77)	5.5 (0.92)	0.19% (1.36)

Note. Standard deviations are shown in parentheses. Word frequency was measured from Francis and Kučera (1982).

of 4.0, while the late OUP words had an average OUP of 6.8. The early and late OUP words were closely matched on the frequency of the initial trigram (24.8 and 26.1 counts in the corpus taken from Mayzner, Tresselt, & Wolin, 1965), overall word frequency (7.7 and 8.0 counts per million taken from Francis & Kučera, 1982), word length (7.4 and 7.6 letters) and the number of syllables (2.5 and 2.5). For each pair, a sentence was created in which the early or late OUP word would fit naturally into the sentence context. Three norming studies were conducted to verify that the early and late words did not differ on three dimensions: participants' familiarity with the target words, goodness of fit into the sentence context and the word's predictability in the sentence frame. Norming data were obtained for the target words and the sentences from separate groups of participants (none of whom participated in the actual experiment).

The first norming study involved 32 participants rating the familiarity of target words on a 1-7 scale. Participants were instructed to provide lower ratings for words that they did not know the meaning of and that they did not use and high ratings for words that they knew the meaning of and used frequently. Due to the length of the list of words to be rated, we constructed three versions of the norms and each participant completed only one version. From this list of words, we selected early and late OUP words, which did not differ in their familiarity ratings (5.8 and 5.7, respectively, see also Table 1).

The second norming study involved 26 participants rating how well a particular target word fit into a sentence context on a 1-7 scale (again with higher values indicating a better fit). In this norming study, participants were instructed to rate words that fit naturally into the sentence context with a higher rating and to provide a lower rating for those words that did not fit naturally into the sentence context. In order to verify that participants were following instructions, we included trials where the target word was either ungrammatical and/or implausible in the sentence context. Two versions of the questionnaire were constructed so that each participant only rated each target word and sentence frame once. Participants rated the early and late OUP words as fitting equally well into the sentence context (5.6 and 5.5, respectively).

Finally, the predictability of the target word from the preceding context was rated by another group of 24 participants by asking them to provide a possible word continuation for the sentence beginning. Participants generated the early and late OUP words less than 1% of the time (only 0.37% for early OUP and 0.19% for late OUP). All of the normative data is presented in Table 1. The only significant difference between the early and late OUP conditions was for the average OUP difference ($p < .001$, all other $ps > .4$)

The boundary paradigm (Rayner, 1975) was used to manipulate the amount of information available from the target word prior to the reader fixating it. The reader was either given a full preview (e.g. *giraffe*), a partial preview (e.g. *girazwp*) or the preview was denied (e.g. *thgjzwp*). When readers made a saccade to the target word, the preview was replaced by the target word (see Fig. 1 for examples). In the full preview condition, the target word itself was the preview; in the partial preview condition, the first four letters were identical to the target and visually dissimilar letters (Buoma, 1973) replaced the remaining letters; and in the preview denied condition, all letters from the target word were replaced with visually dissimilar letters and hence there was no positionally consistent letter information available to the reader in the parafovea. Note that for the early OUP words, the partial preview condition provides consistent letter information up through the OUP, but the same is not true for the late OUP words. Each participant only saw one word from each pair and all target words functioned as nouns in the sentence frame. They thus saw 54 sentences, half containing the early OUP word and half containing the late OUP words. A third of the sentences appeared in each of the three preview conditions, in a fully counterbalanced design. These 54 sentences were included with 60 filler sentences. These filler sentences served as materials in another non-related study and also included a display change. The materials were similar in kind to the experimental materials reported here.

Results and discussion

Two sets of 2 (OUP) × 3 (preview) ANOVAs are reported for three primary measures: first fixation duration (FFD), single fixation duration (SFD) and gaze duration (GD). FFD is the duration of the initial, first pass fixation on a word, SFD is the duration of the fixation when one and only one first pass fixation is made on the target word and GD is the sum of all first pass fixations on a word prior to moving to another word. These measures typically are believed to reflect lexical processing (Rayner, 1998; Rayner & Pollatsek, 1987) or a combination of lexical and integrative processing in the case of gaze duration (Inhoff, 1984) and will be emphasized because they reflect only the first pass reading times on the target word.

Additionally, we analysed two measures: GoPast (GP) and initial landing postion (LP). GP corresponds to the sum of all fixation durations from when readers first enter a region (the target word) until they proceed beyond the target region to the right. Therefore, this measure includes not only first pass fixations, but any time spent rereading earlier text before they proceed on past the target word. This measure reflects a combination of lexical and integrative processing. Finally, LP refers to the letter position in a word where the initial first pass fixation lands (McConkie, Kerr, Reddix, & Zola, 1988; Rayner, 1979).

Trials were excluded from analyses if there was a track loss, if the display change occurred at an inappropriate time or if the reader prematurely ended a trial; this resulted in 18% of the data being lost.[1] Additionally, fixation durations above or below 3 standard deviations from the mean were excluded resulting in an additional 1.5% data loss. Fixations on adjacent letters (when one of the fixations was less than 100 ms) were pooled and individual fixations less than 100 ms and greater than 800 ms were excluded from the analyses.

[1] *Individual participants were excluded if the number of trials excluded exceeded 30%. In Experiment 1, there were 12 participants excluded for this reason.*

The global analyses, which include the fixation data for all trials, are first reported and then analyses including only cases where the reader made a saccade from a location four or less characters to the left of the target region are reported. The choice of four characters was based on the reasoning that the word identification region generally corresponds to approximately eight characters to the right of fixation. This ensures that at the farthest saccade location, four characters from the end of the word, the initial four letters of the target word could fall within the word identification region. Note that the initial four letters of the target word corresponded to the parafoveal preview available to readers in the partial preview condition. In all analyses, the target region refers to the target word and the space before the target word.

Global analyses

Across all measures, there was no benefit for early OUP words compared with late OUP words (see Table 2) on FFD (Fs < 1), GD, $F_1(1, 41) = 1.383$, $p = .246$; $F_2(1, 106) = 2.159$, $p = .145$, or GP, $F_1 < 1$; $F_2(1, 106) = 1.758$, $p = .188$.[2] In fact, readers were numerically faster reading the late OUP than the early OUP words, and significantly faster on SFD by participants, $F_1(1, 41) = 5.020, p = .031$; $F_2 < 1$.

Table 2. Fixation time means as a function of OUP and type of parafoveal preview for the global analyses of Experiment 1

OUP	Preview	First fixation	Single fixation	Gaze duration	GoPast	Landing position
Early	Full	270 (41)	288 (50)	339 (83)	349 (66)	2.78 (.82)
Early	Partial	298 (40)	317 (47)	365 (70)	379 (63)	3.22 (.69)
Early	Denied	291 (42)	340 (86)	373 (77)	404 (80)	2.81 (.72)
Late	Full	269 (39)	289 (49)	340 (72)	352 (78)	2.98 (.67)
Late	Partial	290 (47)	309 (58)	352 (59)	375 (63)	3.14 (.72)
Late	Denied	292 (50)	323 (58)	363 (71)	385 (78)	3.03 (.77)

Note. All first fixation durations, single fixation durations and gaze durations are in ms. The standard deviations are in parentheses.

Although the main effect for preview was significant on all measures (all Fs > 3.7, $p < .05$), the pattern of data was not as robust as expected. There was a numerical advantage for the partial preview over the preview denied condition for all measures, but this difference was only significant or marginally significant on SFD, $t_1(41) = 2.989$, $p = .005$; $t_2(106) = 1.644, p = .103$ and GP, $t_1(41) = 1.838, p = .073$; $t_2(107) = 1.982$, $p = .05$. This contrast did not approach significance on FFD or GD (all ts < 1.1). In contrast, the full preview condition was significantly faster than the preview denied condition on all measures (ts > 2.66, $p < .01$) and significantly faster than the partial preview condition on FFD, SFD and GP (ts $> 3.1, p < .01$), but only by participants on

[2] *We also examined spillover to see if a late OUP effect might emerge. Spillover refers to the duration of the first forward fixation after the reader exits or skips the target word. No significant benefit for an early OUP versus a late OUP was found (Fs < 1).*

GD, $t_1(41) = 2.966, p = .005$; $t_2(107) = 1.634, p = .105$. Contrary to predictions of the Shillcock *et al.* (2000) model, the early OUP words were not read faster than the late OUP words on SFD in the preview denied condition, $t_1(41) = 1.467, p = .150$; $t_2 < 1$. In fact, across all measures in the preview denied condition, the late OUP words yielded numerically faster reading times than the early OUP words (all $ts < 1.47$, $ps > .14$, except GP) and marginally faster by participants on GP, $t_1(41) = 1.784$, $p = .082$; $t_2(106) = 1.554$, $p = .123$. Additionally, the interaction between OUP and preview condition never approached significance on any of these measures (all $Fs < 1.2$).

Additionally, we examined the location of the initial landing position as a function of OUP and preview condition. Readers' initial fixation into the target word landed at roughly the same location for the early and late OUP words (2.94 and 3.05 characters, respectively), $F_1(1, 41) = 1.691$, $p = .201$; $F_2(106) = 1.524$, $p = .22$. Interestingly, readers' LP was affected by the preview condition ($Fs > 4.7$, $ps < .05$). Perhaps surprisingly, readers fixated farther into the target word in the partial preview condition than in the full or preview denied conditions ($ts > 2.6$, $ps < .05$). One might argue that these fixations farther into the target word in the partial preview case reflect that, at least on some trials, readers detected the inconsistent letter information beyond the initial four letters. This provides a potential explanation for why we were unable to obtain a significant benefit of partial information over cases where readers received no positionally consistent letter information.

Near launch site analyses

A subset of the data was analysed for cases where the reader made a saccade from a close launch site (four or less characters from the beginning of the target region) to provide a test of whether readers make use of OUP information under conditions where they are more likely to receive significant parafoveal preview of the target word. A second set of 2 (OUP) $\times$ 3 (preview) ANOVAs were conducted on these data (see Table 3). Due to data loss inherent in this approach as compared with the global analyses,[3] participants and items were included in the planned contrasts if and only if their data contributed to the ANOVAs. Also, for the near launch site cases, we focus on FFD, SFD and GD only because these measures reflect reading times on the target word only and should be the most likely to show lexical effects if present.

For the near launch site cases, there was no suggestion of a main effect for OUP on FFD, SFD or GD (all $Fs < 1$). However, before accepting the null hypothesis that readers do not benefit from an early OUP, it is necessary to examine the preview data more closely.

The main effect for preview was highly significant for FFD, SFD and GD (all $Fs > 6.0$, $ps < .01$) and the full preview condition was always read significantly faster than the partial or preview denied condition ($ts > 2.25$, $ps < .05$). In the partial preview condition, reading times did not differ significantly from the preview denied condition ($ts < 1$). In the partial preview condition, readers likely detected the anomalous

---

[3] *For the near launch site analyses, the participants' means were based on data from 3.1 items per participant mean on average and the item means were based on an average of 2.73 participants per item mean across all three measures. In terms of the participants' analyses here, the FFD and GD analyses included a total of 36 participants and SFD included 26 participants.*

Table 3. Fixation time means as a function of OUP and type of parafoveal preview for the near launch site analyses of Experiment 1

OUP	Preview	First fixation	Single fixation	Gaze duration
Early	Full	272 (48)	285 (48)	309 (91)
Early	Partial	320 (48)	326 (55)	352 (74)
Early	Denied	305 (58)	338 (55)	375 (83)
Late	Full	282 (74)	295 (78)	324 (102)
Late	Partial	311 (70)	329 (65)	360 (112)
Late	Denied	317 (48)	340 (48)	370 (80)

Note. All first fixation durations, single fixation durations and gaze durations are in ms. The standard deviations are in parentheses.

information beyond the initial four letters on some portion of the trials perhaps because of the unusual configuration caused by using visually dissimilar replacement letters. This may have caused a processing cost yielding more similar reading times for the partial and preview denied condition. Otherwise, one would expect shorter reading times for the partial preview versus the preview denied condition across all measures. As a result of this potential anomaly, we emphasize the data from the full preview condition, which avoids any potential issues of readers detecting the visually dissimilar information in the parafovea.

For the full preview condition, there was a trend towards an OUP benefit on FFD and GD; however, neither of these trends approached significance (all ts < 1) and are likely spurious. There was also no significant interaction on any of the measures (Fs < 1.38, ps $> .25$).

One question that numerical differences raise is whether with sufficient power one might obtain a benefit for an early OUP word over a later OUP. This is of more concern for the near launch site cases where data loss is greatest. Essentially, the concern is that we failed to reject the null hypothesis that there is no difference between early and late OUP words when in fact, with sufficient power, we would have found a real difference between early and late OUP words (Type II error). However, even if we found a benefit for an early OUP for near launch site cases, it is not sufficient in and of itself to argue for a serial, sequential letter processing account. For this account to be accurate, one should find a benefit for an early OUP across all launch sites, which clearly was not the case. Unfortunately, there remains the question why readers did not benefit from partial preview information as compared with the preview denied condition. It seems likely that readers processed some of the inconsistent letter information, particularly in the near launch site case, which resulted in a processing cost compared with the full preview condition.

EXPERIMENT 2

The results of Experiment 1 suggest that readers do not benefit from a word possessing an early OUP. One could argue though for a more cautious interpretation. First, the parafoveal preview benefit obtained was smaller than one might have expected based on previous research (e.g. Henderson & Ferreira, 1990; Inhoff, 1989;

Inhoff & Rayner, 1986; Lima & Inhoff, 1985; Pollatsek, Rayner, & Balota, 1986; Rayner, Well, Pollatsek, & Bertera, 1982). Second, our interpretation of the results accepted the null hypothesis. Both of these points warrant a replication of the findings in Experiment 1.

The design of Experiment 2 was nearly identical to Experiment 1. Participants read sentences with either early or late OUP words while their eye movements were monitored. Again, three preview conditions existed: full preview, partial preview, and preview denied condition. In Experiment 2, we replaced letters with lowercase xs rather than with visually dissimilar letters as in Experiment 1. This change should reduce the likelihood that the preview manipulation would interfere with parafoveal processing as compared with visually dissimilar letters.[4]

Second, an additional stimulus control was incorporated based upon insight from Lamberts (2005). He noted that the early OUP stimuli used by Kwantes and Mewhort (1999a) shared fewer positionally consistent letters, or overlapped less in general terms, with other words than the late OUP words and argued that parallel letter processing models could be sensitive to the differences in the degree of letter overlap with other words. Simply put, words that overlap less would take less time to name because fewer letters need to be identified to uniquely identify the word. To control for this possible confound, our stimuli were matched, on average, in terms of the number of words that shared four positionally consistent letters with the target word using CELEX (Baayen, Piepenbrock, & Gulikers, 1995). Importantly, the predictions for the results remain identical to Experiment 1.

Method

Participants

Participants in the experiment were 30 University of Massachusetts community members. All other inclusion criterion were identical to Experiment 1.

Apparatus

The apparatus was identical to Experiment 1.

Procedure

The procedure was identical to Experiment 1.

Stimuli

Thirty pairs of early[5] and late OUP words derived from the MRC database corpus (Coltheart, 1981), closely matched on several variables (see Table 4), were embedded in two different sentence frames written so that both the early and late OUP word fit

[4] *In Experiment 1, letters that could be described as ascenders were replaced with descenders and vice versa. This letter replacement strategy could and generally did change the gross featural shape information about the whole word. By using xs, the change in shape of the target word should be reduced as compared to Experiment 1.*

[5] *After completing the second experiment, it became apparent that five early OUP words (*tornado, saloon, glitter, avalanche, chocolate*) and two late OUP words (*apology, plague*) actually had OUP locations at the fifth location according to a standard American English dictionary. Item analyses were performed excluding these items, and the pattern of results remained the same.*

Table 4. Stimuli properties as a function of OUP condition (early vs. late) in Experiment 2 including average OUP, word frequency (freq.), initial trigram frequency (trigram freq.), word length (length), number of syllables (syll.), number of words that share four position consistent letters with the target word (overlap), familiarity (fam.), target word goodness-of-fit (GOF) in the sentence context and predictability from prior sentence context (predict)

	OUP	Freq.	Trigram freq.	Length	Syll.	Overlap	Fam.	GOF	Predict
Early	4.17 (0.4)	9.5 (10.8)	20.8 (33.0)	7.5 (0.97)	2.5 (0.57)	38.1 (37.3)	6.0 (0.57)	5.6 (0.80)	0.05% (0.00)
Late	6.7 (0.8)	13.0 (8.8)	19.6 (27.5)	7.6 (1.22)	2.5 (0.54)	37.9 (36.2)	5.9 (0.63)	5.6 (0.71)	0.11% (0.01)

Note. Standard deviations are shown in parentheses.

equally well. The early OUP words had an average OUP of 4.17, while the late OUP words had an average OUP of 6.7. The early and late OUP words were closely matched on the frequency of the initial trigram (20.8 and 19.6 counts in the corpus taken from Mayzner *et al.*, 1965), overall word frequency (9.5 and 13.0 counts per million taken from Francis & Kučera, 1982), word length (7.5 and 7.6 letters) and the number of syllables (2.5 and 2.5). Additionally, we controlled for the number of letters that a target word shared with other words. To do this, we matched words in terms of the number of words that shared four positionally consistent letters with the target (Baayen *et al.*, 1995). The early OUP words shared four positionally consistent letters with 38.1 words and the late OUP words shared four with 37.9 words. The early and late OUP words did not differ on any of these stimuli dimensions (all ps $> .4$ except frequency $p = .18$ and OUP $p < .001$).

Normative data were collected on the materials including: word familiarity, goodness of fit in the sentence context, and ratings of predictability in sentence context. Word familiarity ratings were taken from familiarity norms collected for Experiment 1. For the items used in Experiment 2, the early and late OUP words did not differ in terms of their rated familiarity (6.0 and 5.9, respectively).

Thirty-two University of Massachusetts community members rated the target words in terms of the goodness of fit in the sentence on a 1–7 scale (with higher numbers indicating a better fit). This norming procedure was repeated twice to obtain a sufficient number of items. Unfortunately, 1 additional item was needed to obtain a full stimuli set. This item was not formally normed and is not included in the GOF ratings in Table 4. Participants read the beginning sentence context up through the target word and then rated how well the target word functioned as the next word in the sentence context. Participants rated the early and late OUP words as fitting equally well into the preceding sentence context (both 5.6).

Predictability was assessed by having a third group of eight students provide a possible word continuation for the sentence beginnings. The early and late OUP words were not predictable from the preceding sentence context. Participants only generated the appropriate target word less than 1% of the time (0.05% for the early and 0.11% of the time for the late OUP words, ps $> .4$ for all norms). Additionally, there were 14 sentence frames that were not included in the initial predictability ratings. A second less formal norming of these items occurred. Here, the participants consisted of six graduate students and University of Massachusetts community members who were simply asked whether they thought that either the early or the late OUP word was predictable from the sentence context. None of these sentences were judged predictable. The normative data in Table 4 does not include data from this less formal norming session.

Prior to fixating each target word, participants saw one of three preview conditions: full preview, partial preview or preview denied. These preview conditions were identical to Experiment 1 with the exception that the replacement letters were xs in the partial and preview denied condition. Each participant saw all target words without repeating sentence frames. They thus saw 60 sentences, half containing an early OUP word and half containing a late OUP word. Additionally, the target word always functioned as nouns in the sentence frame. A third of these were in each of the three preview conditions, in a fully counterbalanced design. These 60 sentences were included with 54 filler sentences. The filler sentences were used as materials for another experiment. The sentences were similar in kind to the experimental materials and included a display change.

Results and discussion

As in Experiment 1, two sets of 2 (OUP) × 3 (preview) ANOVAs were conducted; one on data from all launch sites (global) and the second on data from launch sites that were four characters or less from the target region (near launch sites). The exclusion criteria were identical to Experiment 1 and resulted in 11.5% of the data being eliminated.[6]

Global analyses

Table 5 displays the participants' means for these analyses. As in Experiment 1, readers did not show a significant benefit for an early OUP word compared with a late OUP word across FFD, GD or GP (all *F*s < 1), but rather late OUP words were generally read numerically faster than early OUP words and marginally faster by items on SFD, $F_1 < 1.3, F_2(1, 116) = 2.970, p = .087$.[7] On FFD, the pattern of data was somewhat less consistent. In the partial preview condition, the FFDs for the early OUP words were marginally faster than for the late OUP words (14 ms), by participants only, $t_1(29) = -1.743$, $p = .092$, $t_2(118) = -1.374$, $p = .172$), but note that the opposite effect is present in the full preview condition, $t_1(29) = 2.013$, $p = .053$, $t_2(118) = 1.372$, $p = .173$. The FFD trend towards a benefit for early OUP words disappeared when looking at SFD, GD and GP though (all ts < 1, except GP, $t_1(29) = 1.560, p = .130$, $t_2 < 1$. In contrast, the trend towards faster reading times for the late OUP words in the full preview condition remained across SFD, $t_1 < 1$, $t_2(116) = 1.466$, $p = .145$, and GD, $t_1(29) = 1.803$, $p = .082$, $t_2(118) = 1.204$, $p = .231$.

Table 5. Fixation time means as a function of OUP and type of parafoveal preview for the global analyses of Experiment 2

OUP	Preview	First fixation	Single fixation	Gaze duration	GoPast	Landing position
Early	Full	303 (46)	306 (49)	335 (45)	346 (57)	3.51 (.65)
Early	Partial	322 (47)	347 (47)	376 (45)	393 (48)	3.37 (.69)
Early	Denied	341 (48)	365 (53)	391 (48)	402 (56)	3.51 (.79)
Late	Full	288 (43)	299 (46)	318 (53)	340 (65)	3.74 (.84)
Late	Partial	336 (39)	346 (43)	377 (52)	382 (56)	3.65 (.82)
Late	Denied	342 (49)	358 (50)	396 (56)	405 (60)	3.37 (.72)

Note. All first fixation durations, single fixation durations and gaze durations are in ms. The standard deviations are in parentheses.

Although we found a marginally significant benefit for early OUP words in the partial preview condition on FFD in this experiment, it seems most parsimonious to suspect that this finding is anomalous. First, this result is only marginally significant by participants and not by items and does not occur on any of the other first pass measures analysed here. Second, when the results are taken as a whole, it seems clear that readers

[6] *In Experiment 2, five participants were excluded from analyses because greater than 30% of their data were eliminated due to a combination of track losses, problems with display change timing and the participant prematurely terminating a trial.*
[7] *We also examined spillover to see if a late OUP effect might emerge. No significant benefit for an early OUP versus a late OUP was found (Fs < 1.2).*

did not consistently benefit from a word possessing an early or late OUP, but that rather we are seeing a natural variability in reading times.

In terms of the main effect for preview, these results more closely resembled previously reported results. The main effect for preview was highly significant across FFD, SFD, GD and GP (all Fs > 27, $p < .001$). Critically, in contrast to Experiment 1, readers read words faster when given partial preview information than when preview was denied: FFD, $t_1(29) = 2.289$, $p = .03$, $t_2(119) = 2.040$, $p = .044$; SFD (by participants only) $t_1(29) = 2.522$, $p = .017$, $t_2(117) = 1.513$, $p = .133$; GD, $t_1(29) = 3.100$, $p = .004$, $t_2(119) = 2.530$, $p = .013$; and GP, $t_1(29) = 2.767$, $p = .01$, $t_2(119) = 2.406$, $p = .018$. In addition, the full preview condition yielded shorter reading times across all measures than either the partial preview or preview denied condition (all ts > 5.2, $p < .001$). Preview condition and OUP did interact significantly by participants on FFD, $F_1(1, 29) = 3.165$, $p = .05$, $F_2(2, 236) = 1.703$, $p = .184$, and approached significance by participants on GD, $F_1(1, 29) = 2.352$, $p = .104$, $F_2 < 1$, but not on SFD nor on GP (Fs < 1). The interaction on FFD is due to the different numerical trends for the partial and full preview conditions discussed earlier.

We also examined whether readers initial, first pass fixation on the target word was affected by OUP. Again, a word's OUP did not significantly affect the LP and resulted in only about 0.1 of a character difference, $F_1(1, 29) = 2.279$, $p = .142$, $F_2(1, 118) = 1.084$, $p = .30$. In contrast to Experiment 1, the type of preview did not significantly affect readers' initial landing position in the target word, $F_1(1, 29) = 1.818$, $p = .171$, $F_2(2, 236) = 1.290$, $p = .277$. The pattern of means for the three preview conditions was more in line with what one might expect with the full preview condition yielding fixations farther into the target word than the partial or no preview condition. The interaction between OUP and preview did not reach significance either, $F_1(2, 58) = 1.472$, $p = .238$, $F_2(2, 236) = 2.126$, $p = .122$.

Near launch site analyses

As in Experiment 1, when analysing the near launch site data, we focused on FFD, SFD and GD. When collapsed across preview conditions for the near launch site cases, the reading times for late OUP words were always shorter than for early OUP words across all measures (see Table 6),[8] but did not reach significance for FFD ($F_1 < 1$, $F_2(1, 71) = 2.235$, $p = .139$), and only by items on SFD ($F_1 < 1$, $F_2(1, 64) = 4.385$, $p = .04$). Importantly, reading times were marginally faster for late OUP words than early OUP words on GD, $F_1(1, 27) = 3.912$, $p = .058$, $F_2(1, 70) = 3.776$, $p = .056$. This effect on GD is driven primarily by the difference between the early and late OUP words in the preview denied condition. Although we found marginally faster reading times for late OUP than early OUP words on GD, it is important to examine the direction of the means in each contrast involving OUP to see if there was any suggestion of a benefit for an early OUP given that we are in part arguing to accept the null hypothesis.

There are two contrasts where early OUP words yield numerically shorter reading times than late OUP words: first, for FFD in the partial preview condition, and secondly, for SFD in the full preview condition. The trend towards a benefit for early OUP words

[8] *For the near launch site analyses, the participants' means were based on 3.52 items per participant mean on average and the item means were based on an average of 2.2 participants per item mean across all three measures. In terms of the participants' analyses here, the FFD and GD analyses included a total of 28 participants and SFD included 26 participants.*

Table 6. Fixation time means as a function of OUP and type of parafoveal preview for the near launch site analyses of Experiment 2

OUP	Preview	First fixation	Single fixation	Gaze duration
Early	Full	289 (62)	286 (65)	300 (71)
Early	Partial	342 (48)	356 (53)	384 (63)
Early	Denied	367 (61)	377 (61)	411 (66)
Late	Full	287 (50)	296 (55)	302 (54)
Late	Partial	352 (47)	357 (49)	372 (62)
Late	Denied	351 (56)	353 (55)	382 (69)

Note. All first fixation durations, single fixation durations and gaze durations are in ms. The standard deviations are in parentheses.

was not stable ($ts < 1.15$, $p > .25$) and only occurred for FFD for the partial preview condition and disappeared for SFD and GD consistent with the trend in the global data. The same is true for the full preview condition for SFD. The trend was neither stable ($ts < 1.15$, $p > .30$) nor consistent when comparing these times to GD. Essentially, when readers made a saccade from near launch sites, they did not benefit from an early OUP word.

Given the lack of a clear benefit for an early OUP word, it is important to examine whether the standard preview effect was present. Again, the main effect of preview condition was highly significant on all measures (all $Fs > 18$). That is, participants read words in the full preview condition faster than in the partial or preview denied conditions (all $ts > 4.3$). The partial preview condition yielded shorter reading times than the preview denied condition for FFD (by items), $t_1(27) = 1.516$, $p = .141$, $t_2(72) = 2.297$, $p = .025$, and by participants only for GD, $t_1(27) = 1.941$, $p = .063$, $t_2(71) = 1.252$, $p = .215$. However, it did not reach significance for SFD, $t_1(26) = 1.145, p = .262$, $t_2(65) = 1.322, p = .191$. As in Experiment 1, it seems likely that readers detected the replaced letters (xs) in the partial preview condition which led to longer reading times. Finally, there were no significant interactions present for any of the first pass measures (all $Fs < 1.65$).

GENERAL DISCUSSION

The present results provide no evidence of an advantage for early OUP words in reading. Even when only near launch site cases were examined, which should have optimized the likelihood of finding a benefit for an early OUP, readers showed no significant benefit. As discussed earlier, this conclusion is based on null results, which one might argue could be due to a lack of power in the design. This is especially true in the near launch site cases, as only a subset of the data was considered. In order to provide further evidence against this objection, we pooled the data from Experiments 1 and 2 and recomputed the near launch site analyses using Experiment as a between-participant and between-item variable. While this increased our degrees of freedom dramatically, early OUP words still did not show any advantage over late OUP words. In fact, all numerical differences produced a small advantage for late OUP words. This reverse main effect of OUP only reached significance in gaze duration by items, 18 ms; $F(1, 157) = 4.47$, $p = .036$. In addition, OUP did not significantly interact with any other variable. Hence, the present results provide no evidence supporting serial sequential letter processing during sentence reading.

These results leave two obvious questions unanswered: First, how does one reconcile the present data with that of OUP effects in other tasks (Kwantes & Mewhort, 1999a; Lindell *et al.*, 2003)? Second, what ramifications do these data have for models of reading and visual word recognition? Each of these questions will be addressed below.

The differences between Kwantes and Mewhort (1999a) and our data appear to be due primarily to differences in the experimental methodology used: eyetracking during sentence reading versus naming. In an attempt to confirm this hypothesis, we generated simulated naming times by submitting our target words to the English Lexicon Project's (ELP) database (Balota *et al.*, 2002). The database includes behavioural information about all target words minus one from Experiment 2. For Experiments 1 and 2, the items yielded numerically faster naming times for early OUP than late OUP words; the effect size was 17 ms for Experiment 1 and 1 ms for Experiment 2, while Kwantes and Mewhort's original stimuli (excluding 1 item that was not in the database) yielded an OUP effect size of 19 ms using the ELP database. The size of the effect is smaller than the 29 ms effect reported in the Kwantes and Mewhort paper. This difference in the size of the effect could be explained by considering that Kwantes and Mewhort blocked the presentation of their stimuli to maximize any OUP effect, which was not the case in the ELP database.

In a lexical decision task, Lindell *et al.* (2003; Experiment 1) reported a 33 ms benefit for early OUP words. When submitted to the ELP database (Balota *et al.*, 2002), Lindell *et al.*'s items only produced a 10 ms effect. This large reduction in the size of the effect could be due to Lindell *et al.*'s use of stimuli repetition in their design; in their study, each item was repeated 3 times, which differs from the ELP database. Finally, Lindell *et al.* used non-words with an orthographic deviation point from words that was either early or late. The nature of the non-words in the task could have exaggerated the size of the OUP effect. Interestingly, when naming times from the ELP database are generated for Lindell *et al.*'s words, there is a large 33 ms advantage for early OUP words over late OUP words. When the stimulus sets from Experiment 1 were submitted to the ELP database, they also produced a smaller OUP effect on lexical decision times (14 ms) compared with the word naming results presented above. In addition, the stimuli from Experiment 2 (where letter overlap between early and late words was controlled) produced a reverse effect of OUP in lexical decision times obtained from the ELP database, with late OUP words on average being responded to 12 ms faster than early OUP words. This reversal of the effect suggests that the OUP effect in lexical decision is not consistent. In fact, when lexical decision times for the Kwantes and Mewhort (1999a) stimuli set are generated from the ELP database, these items also show a 9 ms reverse OUP effect.

Taken together, the results from the current studies and the simulations from the ELP database (Balota *et al.*, 2002) present an interesting picture of the nature of OUP effects. OUP effects appear to be consistently observed in word naming across three different stimulus sets, where letter overlap with other words is not controlled. The original Kwantes and Mewhort (1999a) stimuli set (19 ms) and the Experiment 1 stimuli set (17 ms) produce comparable results in naming. This provides evidence that the failure to find an OUP effect in Experiment 1 is not simply due to the stimuli selected. The stimulus set from Experiment 2 shows no benefit for early OUP words (1 ms) in word naming. This elimination of the OUP effect is most probably due to the extra stimulus control used in this set (see discussion below). The ELP simulations also demonstrate that OUP effects are not as consistent in the lexical decision task. The simulations produced a 10 ms effect for Lindell *et al.*'s (2003) stimulus set, a 14 ms effect for the Experiment 1 stimulus set, a 12 ms advantage for late OUP words using the Experiment 2 stimulus set and a 9 ms advantage for the late OUP words from the original Kwantes and Mewhort stimulus set. Finally, the experiments reported in this paper provide no evidence for an OUP effect during sentence reading.

The best test of whether 'true' OUP effects are observed in various tasks would be to consider only the stimulus set from Experiment 2. In this stimulus set, an additional variable was controlled, overlap between early and late OUP words in terms of other words in the lexicon. This control was added to account for Lamberts' (2005) critique of Kwantes and Mewhort's (1999a) stimuli. Lamberts pointed out that early OUP words used by Kwantes and Mewhort overlapped with fewer other words in the English language compared with late OUP words. Therefore, the stimulus set from Experiment 2 allowed a test of whether a 'true' OUP effect is observed over and above its confound with letter overlap. With the stimulus set of Experiment 2, there was no advantage for early OUP words in naming or in the lexical decision task (using the ELP database). The preview denied condition in Experiment 2 provides the closest analogue to an isolated word presentation (since there was no useful parafoveal preview except for the length of the target prior to its first fixation in this condition) and produced no significant benefits for an early OUP. Thus, these results suggest that 'true' OUP effects do not exist in any task.

The discrepancy between eyetracking and word naming when overlap of letters is not controlled is rather surprising because it is generally assumed that they have common underlying mechanisms that tend to be correlated (Schilling, Rayner, & Chumbley, 1998). That being said, it is not unreasonable to expect that part of the difference in findings between our data set and that of Kwantes and Mewhort (1999a) could be attributable to task differences. The Schilling *et al.* paper involved a word frequency manipulation which is quite robust across experimental tasks. Other phenomena may produce a divergence in findings between tasks (see Murray & Forster, 2004, for a discussion). OUP is such a phenomenon, with it only exerting a consistent effect in word naming when letter overlap is not controlled. One obvious locus of the OUP effect in naming would be in the articulation of words. However, this hypothesis was effectively ruled out by Kwantes and Mewhort who failed to show an effect of OUP in a delayed naming task, where identification of the word and access to its phonology is assumed to occur prior to the cue to name the word. Instead, some aspect specific to the word naming task, other than articulation itself, must be responsible for the OUP effect observed in this task.

The lack of a generalizable OUP effect poses a problem for all models of word recognition that posit letters are identified for word recognition or used to retrieve lexical information in a serial fashion. One such model is LEX (Kwantes & Mewhort 1999b), which has been used to simulate a variety of phenomena in word naming and lexical decision data. This model consists of three parts: letter identification, retrieval of the word from the lexicon and generation of response. As noted in the Introduction, letter identification is accomplished in parallel. However, in order to retrieve the word from the lexicon, the identified letters are stored in an ordered list (based on the left-to-right ordering of the letters in the word) and this list is used to probe memory sequentially. As more letters are used to probe memory, the number of possible candidate words in the lexicon reduces. This reduction in candidates should thus result in shorter lexical identification for early OUP words. Both orthographic and phonological information are retrieved from memory simultaneously. However, the phonological representation needs to be 'deblurred' in order to name a word, so searches of the lexicon are orthographically based in both word naming and lexical decision.

For word naming, the information in the phonological representation is used to generate a response. Response latency is simulated as the number of times memory was probed prior to identifying the word, plus the time taken to 'deblur' the phonological representation (which takes longer for words with spelling-to-sound irregularities). Lexical decision reaction time is simulated by the number of times memory was probed plus a random-walk component. If the non-word/word decision is particularly difficult, a verification process will be enacted that checks the phonological representation of the word prior to a decision regarding its lexical status.

The important factor in this model is that the OUP effect occurs during the retrieval process itself, due to the fewer number of times that memory must be probed for early OUP words compared with words with late OUP words. Although Kwantes and Mewhort (1999b) did not actually simulate OUP effects with LEX, it should predict consistent effects of OUP in both tasks, as the retrieval mechanism is the same. As illustrated above through the use of the ELP database (Balota *et al.*, 2002), OUP effects are not consistent in the lexical decision task. It appears that the model could possibly predict a smaller (or reverse) effect of OUP in lexical decision if a bias is added to the random-walk component in lexical decision in favour of late acquired words. However, there is no theoretical reason for biasing the lexical decision response in this way. Moreover, while LEX is not a model of word recognition during sentence reading, it is not parsimonious to assume that retrieval of a word from memory would be accomplished in a different fashion in tasks such as naming and lexical decision compared with sentence reading. Thus, it appears that the serial word retrieval component of LEX is simply not plausible given the current data.

As mentioned in the Introduction, the results from the current experiment also have implications for the split processing model (Shillcock *et al.*, 2000). This model claims to take into account the anatomy of the fovea, which Shillcock *et al.* claim is split into two sections, each projecting to the primary visual cortex in the contralateral hemisphere. They argue that during reading, words are split at the fixation point and that these two parts of the words are acted upon separately at an early stage of processing during word recognition (i.e. that the image is not fused prior to the word recognition process). The hemispheres each act upon their word part to identify the word by finding words in the lexicon that match the letters in the part, and this information is transmitted across hemispheres. Shillcock *et al.* also provide evidence through a corpus analysis that the optimal position for the split, where each hemisphere has an equal chance of identifying the word, is either to the centre or slightly to the left of centre for isolated words, a result that is consistent with the optimal viewing position observed in isolated word recognition experiments (O'Regan & Lévy-Schoen, 1987).

Following the assumptions made in the introduction, the split processing model should predict shorter reading times for early versus late OUP words unless one assumes precise coding of length, which is unlikely, particularly for longer words (Fischer, 2000). The split processing eye movement model is a model of isolated word recognition and our preview denied condition provides the closest analogue to this type of presentation because the reader obtains no useful preview of the target word other than length information. From our data, particularly that of Experiment 2, the predicted OUP effect was clearly not present.

In summary, the lack of benefit for early OUP words in reading, even when conditions were optimized such that readers should benefit from making use of this information, as in the near launch site cases, provides evidence against serial, sequential letter processing during reading. The difference in results between the Kwantes and Mewhort (1999a) experiment and our Experiment 1 appears to be due to task demands specific to word naming. However, our results from Experiment 2 combined with the simulated naming and lexical decision data generated for these target words bring into question whether a true OUP effect exists for any task when letter overlap is controlled (Lamberts, 2005). The present findings shed doubt on the validity of models that would predict a general effect of OUP in word recognition, such as LEX (Kwantes & Mewhort, 1999b). This finding also appears to present difficulties for the split processing model of Shillcock *et al.* (2000).

Acknowledgements

This research was supported by Grant HD26765 from the National Institute of Health. Experiment 2 was conducted as part of the first author's doctoral dissertation. During the preparation of this article, the first author was supported by a post-doctoral fellowship on Grant HD-07548 from the National Institute of Health and the second author was supported on a pre-doctoral fellowship on Grant MH16745 from the National Institute of Mental Health. The data were presented in part at the Annual Meeting of the Psychonomic Society in November, 2003. We would like to thank Guo Li Yan and Debra Machacek for their help with development of the stimuli and collection of the normative data, and Dave Braze, Simon Liversedge, Wayne Murray, Donald Shankweiler, Geoff Underwood and an anonymous reviewer for their helpful comments on an earlier draft.

References

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

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*. Retrieved from http://elexicon.wustl.edu/

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language*, *48*, 615-634.

Brysbaert, M. (1994). Interhemispheric transfer and the processing of foveally presented stimuli. *Behavioural Brain Research*, *64*, 151-161.

Buoma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, *13*, 767-782.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, *33A*, 497-505.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed processing approaches. *Psychological Review*, *100*, 589-608.

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.

Fischer, M. H. (2000). Word centre is misperceived. *Perception*, *29*, 337-354.

Francis, W., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Henderson, J., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 417-429.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, and Computers*, *16*, 96-101.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1612-1627.

Inhoff, A. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior*, *23*, 612-624.

Inhoff, A. (1989). Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 544-555.

Inhoff, A., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics*, *40*, 431-439.

Kwantes, P., & Mewhort, D. (1999a). Evidence for sequential processing in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 376-381.

Kwantes, P., & Mewhort, D. (1999b). Modeling lexical decision and word naming as a retrieval process. *Canadian Journal of Experimental Psychology*, *53*(4), 306-315.

Lamberts, K. (2005). Left-to-right sequential processing in visual word recognition? Alternative interpretation of orthographic uniqueness point effects. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 14-19.

Lima, S., & Inhoff, A. (1985). Lexical access during eye fixations in reading: Effects of word-initial letter sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 272-285.

Lindell, A., Nicholls, M., & Castles, A. E. (2003). The effect of orthographic uniqueness and deviation points on lexical decisions: Evidence from unilateral and bilateral-redundant presentations. *Quarterly Journal of Experimental Psychology*, *56A*, 287-307.

Mayzner, M., Tresselt, M., & Wolin, B. (1965). Tables of trigram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements*, *1*, 33-78.

McConkie, G., Kerr, P., Reddix, M., & Zola, D. (1988). Eye movement control during reading: I. The location of the initial eye fixations in words. *Vision Research*, *28*, 1107-1118.

McConkie, G., & Zola, D. (1987). Visual attention during eye fixations while reading. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 385-401). London: Erlbaum.

Murray, W., & Forster, K. (2004). Serial mechanisms in lexical access: The Rank hypothesis. *Psychological Review*, *11*, 721-756.

O'Regan, J., & Lévy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance: Vol. 12. The psychology of reading* (pp. 363-383). Hillsdale, NJ: Erlbaum.

Pacht, J. M. (2003, August). *Uniqueness point effects in visual word recognition.* Poster session presented at the 12th European Conference on Eye Movements, Dundee, UK.

Plaut, D., McClelland, J., Seidenberg, M., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56-115.

Pollatsek, A., Rayner, K., & Balota, D. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception and Psychophysics*, *40*, 123-130.

Pynte, J., Kennedy, A., & Murray, W. (1991). Within-word inspection strategies in continuous reading: Time course of perceptual, lexical, and contextual processes. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 458-470.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*, 65-81.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception*, *8*, 21-30.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12*, 206-226.
Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance* (Vol. 12, pp. 327-362). London: Erlbaum.
Rayner, K., Well, A., Pollatsek, A., & Bertera, J. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics, 31*, 537-550.
Reichle, E. D., Pollatsek, A., Fisher, D., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*, 125-157.
Reichle, E. D., Pollatsek, A., & Rayner, K. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445-526.
Roberts, M., Rastle, K., Coltheart, M., & Besner, D. (2003). When parallel processing in visual word recognition is not enough: New evidence from naming. *Psychonomic Bulletin and Review, 10*, 405-414.
Schilling, H., Rayner, K., & Chumbley, J. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory and Cognition, 26*, 1270-1281.
Seidenberg, M., & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.
Shillcock, R., Ellison, T., & Monaghan, P. (2000). Eye-fixation behavior, lexical storage, and visual word recognition in a split processing model. *Psychological Review, 108*, 824-851.
Underwood, N., & McConkie, G. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly, 20*, 153-162.

Received 29 September 2004; revised version received 28 July 2005

Appendix A

Stimuli for experiment 1

(1) My mother's loud (beagle, collie) barked at the mailman.
(2) Jane saw the famous (athlete, princess) walking down the street.
(3) Jill added some chopped (banana, salami) to the dish she was making.
(4) Bob could see the (horizon, factory) from his balcony.
(5) I heard that John's (intern, butler) just quit yesterday.
(6) Jen bought some expensive (chocolate, champagne) from a shop in Paris.
(7) Kelly added some fresh (avocado, pepper) to the dish she was making.
(8) The authorities did not know what triggered the (avalanche, dynamite) that killed 10 people.
(9) Samantha received a long (document, bulletin) attached to an e-mail.
(10) Andy went to the new (arcade, casino) where he spent lots of money.
(11) I saw a movie about a large (dinosaur, monster) that invaded Los Angeles.
(12) We stopped at the large (canyon, marina) when we needed to rest.
(13) Brett could see a large (beaver, rooster) right outside his window.
(14) The doctor concluded that Ty's (dyslexia, fracture) was correctly diagnosed.
(15) We bought our son a new (cradle, balloon) when we went to the store.
(16) The ranger said that an adult (dolphin, salmon) can swim hundreds of miles.
(17) Beth bought some (hosiery, lingerie) while she was at the department store.
(18) My teacher told us that a young (gorilla, primate) learned sign language.
(19) Martha prepared a tasty (dessert, pastry) before her dinner guest arrived.
(20) This is the worst (drought, deficit) that the nation has faced in years.

(21) At Macy's, you can buy everything from a basic (blouse, skillet) to a sofa for your house.
(22) Jeanette saw a baby (giraffe, peacock) for the first time at the zoo.
(23) Lisa bought a baby (gerbil, rabbit) from the local pet store.
(24) In the distance, a large (jaguar, panther) is hiding amongst the bushes.
(25) The elderly (janitor, stylist) was suddenly fired from his job.
(26) The reduced price of the expensive (limousine, bracelet) made it more marketable.
(27) The group was distracted by the loud (eruption, diversion) and didn't hear the guide.
(28) Would you run out and buy some (cinnamon, pepperoni) so that I can finish dinner?
(29) I had forgotten what time Jane's (wedding, concert) is supposed to start.
(30) During his trip, John saw a beautiful (gondola, convent) on the banks of a canal in Italy.
(31) The pictures of birds included a young (pelican, cockatoo) being fed by its mother.
(32) Hearing loss kept Sam from playing his beloved (tambourine, accordion) in his local band.
(33) The reporter said that the last (pavilion, sanction) was put in place last month.
(34) In history class, we read about a Scottish (rivalry, ministry), which really intrigued me.
(35) Before leaving the factory, every (moccasin, roadster) is examined by quality control.
(36) The photo showed the main (witness, terminal) collapsing to the ground.
(37) The researcher realized the earlier (trajectory, paradigm) he was using was not accurate.
(38) Jerry already pointed out the (omission, corridor) to Catherine earlier today.
(39) The news clip showed the crook's (silhouette, accomplice) as he was being led away by police
(40) The company sells (kerosene, pesticide) to distributors across the country.
(41) Before the clean-up, the remains of the (scaffold, grenade) were scattered across the ground.
(42) The young woman's (toddler, relative) just got out of the hospital yesterday.
(43) The description of the large (vulture, brothel) brought the scene to life.
(44) Ignoring the only remaining (sceptic, constraint) could end up leading to problems.
(45) What is the local (monument, specialty) in this area called?
(46) My father told me to be careful not to touch the (raccoon, portrait) with my fingers.
(47) After an hour of driving, we found the historic (junction, sanctuary) listed in the book.
(48) Unfortunately, the small (ligament, bracket) is damaged beyond repair.
(49) The decision to remove the last (segment, diplomat) was final and could not be changed.
(50) We ate a lot at the (banquet, barbecue) hosted by my father's company.
(51) I read a story about an evil (gremlin, villain) that terrorized a small town.
(52) I heard that the young (actress, hostess) really loves her job.
(53) After digging, there was no place the young (scavenger, commando) could wash off.
(54) Bob drank a lot of the (beverage, lemonade) because he was terribly thirsty.

Appendix B

Stimuli for experiment 2

(1) The disease took the life of a healthy (athlete, infant) yesterday evening at the clinic.
(2) The security guard led us to the secret (auction, corridor) and told us to go inside.
(3) In 1998, my home-town had the worst (avalanche, diplomat) that it has ever had.
(4) I had heard about the role of the farmer's (boycott, almanac) during pre-industrial times.
(5) Jaime almost walked right into a dead (cactus, rabbit) on her hike through the desert.
(6) At *Big Y*, my wife always tries to sneak another (chocolate, vegetable) into my shopping cart.
(7) I was startled by the loud (dalmatian, diversion) and looked towards it.
(8) After lunch, Sara realized she left her boss's (dessert, bulletin) at the deli next door.
(9) Teresa wanted an immediate (divorce, apology) after the last fight with her husband.
(10) At the auction, the historic (document, portrait) sold for almost a thousand dollars.
(11) John searched for info on the Web about (dyslexia, pneumonia) as part of a class project.
(12) The human toll from the damaging (eruption, sanction) may never be known.
(13) In the gun battle, the soldier shot the fleeing (fugitive, civilian) in the ensuing panic.
(14) Carrie promised to stop and get some (glitter, gasoline) on her way to the party.
(15) Sid decided to order (haddock, whiskey) but the waiter said they were out of it.
(16) After dusk, driving towards the dark (horizon, factory) gives Joan an eerie feeling.
(17) Jay's son seemed to enjoy the black (licorice, panther) more than anything else.
(18) A student asked a question about a particular (ligament, primate) discussed in the reading.
(19) An expert was asked to examine the unknown (mixture, sculpture) that was just found.
(20) After the war, the town's people removed the only (nobility, convent) that was still left.
(21) The dilapidated apartment has one remaining (occupant, balcony) that is still left there.
(22) Harry had the unpleasant job of burying the dead (raccoon, monkey) at the pet cemetery.
(23) Joe found his drunk father at the dirty (saloon, marina) and took him home.
(24) After testing the remaining (sediment, transistor), the technician called it a day.
(25) The plumber repaired the problem with the (sewage, toilet) before the work-day began.
(26) All I could see was a large (silhouette, sanctuary) off in the distance.
(27) The newspaper had an article on another (suicide, veteran) in the local section today.
(28) Jane almost dropped her sister's (toddler, packet) when trying to open the door.
(29) Jeremy's mother has tried to instill (tolerance, discipline) into her young son.
(30) Hundreds were killed by the devastating (tornado, plague) that ravaged the region.

(31) The doctor was investigated after a healthy (athlete, infant) died while under his care.
(32) Walter found the entrance to the private (auction, corridor) after searching for an hour.
(33) The officials mentioned that the last (avalanche, diplomat) was completely harmless.
(34) Sharon said the new (boycott, almanac) starts on the first of the year.
(35) Sara gave her favorite (cactus, rabbit) away to a friend when she moved to London.
(36) Sam decided to eat the (chocolate, vegetable) now rather than later.
(37) The child got distracted by the small (dalmatian, diversion) and lost track of his mother.
(38) Anne was cleaning and picked up the (dessert, bulletin) from the coffee table.
(39) The movie star's public (divorce, apology) appeared in all of the tabloids.
(40) In a fit of rage, Tony destroyed a rare (document, portrait) dating back to the Civil War.
(41) Paul's case of mild (dyslexia, pneumonia) makes it hard for him to complete his homework.
(42) The residents prepared for an even larger (eruption, sanction) by hording fuel oil.
(43) The medic tried to help the wounded (fugitive, civilian) who was bleeding profusely.
(44) Jeff spilled the remaining (glitter, gasoline) that was left in the container.
(45) Sue asked if I had ever tried (haddock, whiskey) and I told her that I never have.
(46) You could see the smoke off in the far-off (horizon, factory) but not the fire.
(47) The teacher asked what type of (licorice, panther) the class liked the best.
(48) The veterinarian operated on the injured (ligament, primate) to prevent further injury.
(49) George examined the concrete (mixture, sculpture) outside very closely.
(50) In a story, the peasants resented the wealthy (nobility, convent) that resided in Paris.
(51) The house's original (occupant, balcony) dates back to the early 1800s.
(52) Unfortunately, the injured (raccoon, monkey) had to be euthanized.
(53) The director imagined that the abandoned (saloon, marina) would be a great backdrop.
(54) I was unclear what to do with any (sediment, transistor) that needed to be
(58) Jim's wife had planned to pick up their (toddler, packet) at day care but she forgot.
(59) Pam felt it was important to exhibit (tolerance, discipline) in her work as a social worker.
(60) The amount of damage that a long-lasting (tornado, plague) causes can be tremendous.