On Coding the Position of Letters in Words
A Test of Two Models

Carol Whitney,1 Daisy Bertrand,2 and Jonathan Grainger3

1Department of Human Development, University of Maryland, College Park, MD, USA, 2Aix-Marseille University, Marseille, France, 3CNRS and Aix-Marseille University, Marseille, France

Abstract. Open-bigram and spatial-coding schemes provide different accounts of how letter position is encoded by the brain during visual word recognition. Open-bigram coding involves an explicit representation of order based on letter pairs, while spatial coding involves a comparison function operating over representations of individual letters. We identify a set of priming conditions (subset primes and reversed interior primes) for which the two types of coding schemes give opposing predictions, hence providing the opportunity for strong scientific inference. Experimental results are consistent with the open-bigram account, and inconsistent with the spatial-coding scheme.

Keywords: visual word recognition, orthographic processing, letter position coding

The last decade has seen an explosion of interest in the study of orthographic processing. Indeed, the study of how the brain encodes letter strings is potentially of great theoretical and practical significance. Understanding how the brain accomplishes this task may shed light on basic processes of visual analysis, and lead to insights into the etiology and remediation of developmental dyslexia (Whitney & Cornelissen, 2005).

As discussed by Grainger and Whitney (2004), it is now evident that a position-specific slot encoding of letter order, such as employed by the Interactive Activation model (McClelland & Rumelhart, 1981), is inconsistent with experimental results (e.g., Davis & Bowers, 2004; Grainger, Granier, Farioli, Van Assche, & van Heuven, 2006; Perea & Lupker, 2003, 2004; Peressotti & Grainger, 1999; Schoonbaert & Grainger, 2004; see Grainger, 2008, for review). Several alternative models have converged on the proposal that lexical access employs representations of ordered letter pairs that are not necessarily contiguous, known as open-bigrams (SERIOL model: Whitney, 2001; OB model: Grainger & van Heuven, 2003; LCD model: Dehaene, Cohen, Sigman, & Vinckier, 2005). For example, the stimulus BIRD would activate open-bigrams BI, IR, RD, BR, ID, and BD (and edge bigrams #B and D#, depending on the model). In contrast, other recent models propose lexical access based on individual-letter representations (SOLAR spatial-coding model: Davis, 1999, 2010; Davis & Bowers, 2006; Overlap model: Gomez, Ratcliff, & Perea, 2008; Bayesian Reader: Norris, Kinoshita, & van Casteren, 2010).

The open-bigram and spatial-coding models, in particular, yield very similar predictions and are difficult to differentiate experimentally. The focus of the present article is to present predictions that differ between these two types of models, and to experimentally evaluate these predictions following the principles of strong inference (Platt, 1964). We first review these input coding mechanisms and discuss how they explain the basic experimental results. We then consider how to distinguish between the open-bigram and spatial-coding models.

The majority of evidence against a position-specific letter encoding comes from masked priming experiments (Grainger & Whitney, 2004). Therefore, any model of orthographic processing must be consistent with priming results. In evaluating a model, it is necessary to provide a metric that specifies how much input a given stimulus (i.e., a prime) provides to the target word—a so-called match value. The ordering of modeled target inputs for different primes should be consistent with amounts of experimentally observed facilitation. Note that it is the ordering of priming effects sizes and not the absolute size of any given priming effect that can be used to distinguish between models. Of course, an even stronger test can be provided on the rare occasions that different models predict effects in the opposite direction, which will be the case in the present study.

In models that implement open-bigram (OB) coding, the calculation of match values is accomplished by specifying the activation pattern of OBs, and the connection weights between OBs and word nodes. A word-node’s activation is then the dot product of the OB vector and the weight vector. That is, each OB’s activation is multiplied by the corresponding weight, and lexical input is the sum of these products.

In its simplest version, OB coding assumes a limit of two intervening letters. For example, the stimulus CARPET would activate open-bigram CP, but not CE. In a binary
OB model, all activations and weights are 1 or 0. In a graded model, activations and weights decrease with increasing separation between constituent letters. OBs activated by contiguous letters reach 1.0; a one-letter separation yields an activation < 1.0, denoted B1; a two-letter separation yields an activation ≤ B1, denoted B2. Typical values of B1 and B2 are 0.8 and 0.4, respectively.

Weights on OB-to-word connections are equal to the OB activation pattern for the given word. For example, under the above parameters for a graded model, the weights between BI, IR, RD, and the word-node BIRD are 1.0, the weights between BR, ID, and BIRD are 0.8, and the weight between BD and BIRD is 0.4. The stimulus BIRD would yield activation of the word-node BIRD equal to

\[
3 \times (1.0 \times 1.0) + 2 \times (0.8 \times 0.8) + (0.4 \times 0.4) = 4.44.
\]

In the spatial-coding scheme implemented in the SOLAR model (Davis, 1999, 2010), bottom-up lexical input is given by a complex function, in which degree of alignment of shared letters is evaluated. See Davis (2010) for details.

Essentially, the spatial-coding scheme computes how well the stimulus preserves the spacing of letters shared by the stimulus and the target word. We can approximate the spatial-coding input function by aligning the prime with the target such that the number of exact matches is maximized, and weighting each shared letter by how far it is shifted. Each exact match contributes 1.0; each letter shifted by one position contributes SI < 1.0; each letter shifted by two positions contributes S2 < SI, while letters off by three or more positions contribute nothing. The current parameters yield values of ~ 0.7 for SI, and ~ 0.2 for S2. Although this is only an approximation, it provides a reasonable computation of similarity under the spatial-coding model. (For simplicity, we ignore possible higher weightings for external letters, which are not relevant given the priming conditions that were tested.) See Figure 1 for examples.

To further illustrate the models, we next review basic priming phenomena and consider how the models account for these results. In the following, we use the standard notation for specifying primes, in which the prime’s letters are represented by their position in the target word, with “d” denoting a letter not in the target. For example, the prime shop for the target SOAP is represented 1d24, because S is the first letter of the target, H is not in the target, O is the second letter of the target, and P is the fourth letter of the target.

Studies have shown that letter order is highly important when the prime is comprised of a restricted subset of the target’s letters. For example, for seven-letter targets, the prime 1357 provides facilitation relative to the control dddd, whereas 1537 does not (Grainger et al., 2006). When all letters of the target are present in the prime, maintenance of letter order is less important, and transpositions then yield an advantage over replacements. For example, for five-letter targets, the prime 12435 provides facilitation relative to a double substitution control prime 12dd5 (Perea & Lupker, 2003). Thus a model of orthographic encoding must offer sufficient sensitivity to letter order to explain the results for 1357 versus 1537, yet yield enough flexibility to explain the results for 12435 versus 12dd5.

It is straightforward to explain these findings under an OB coding scheme. For simplicity, we consider a binary scheme without edge bigrams. For prime 1357, three OBs (13, 35, and 57) match the target; for prime 1537, only two OBs (13 and 57) match the target. Thus 1357 provides stronger activation of the target. Prime 12435 matches all OBs of the target except for 34, while 12dd5 matches only 12 and 25. Thus 12435 provides stronger activation of the target.

Under the spatial-coding model, 12435 provides greater activation of the target than 12dd5, due to partial input generated by letters 4 and 3, which are each shifted by one position. Hence, 12435 yields 2 × SI more input than 12dd5. As shown in Figure 2, 1357 provides stronger activation of the target than 1537 provided that 2SI + S2 > 1.0, because 1537 yields two exact matches, while 1357 gives one exact match, two SIs, and one S2.

![Figure 2](image-url)  
**Figure 2.** Comparison of the amount of overlap between primes 1357 and 1537 for the target 1234567 under the spatial-coding model.

For many prime contrasts, OB and the spatial-coding schemes give similar predictions because changes of letter order reduce the number of matching open-bigrams and disrupt letter spacing. In order to directly compare the models, it is necessary to create primes for which the models inherently give different predictions. As discussed by Whitney (2008), it is possible to match the amount of positional overlap under the spatial-coding model, while varying the degree to which letter order is preserved. The resulting primes yield different predictions under the two types of models. In particular, consider the primes 1237 and 1654327. Both primes match both exterior letters, so we ignore edge effects.

Under an OB scheme, 1237 activates target OBs 12, 23, and 13, while 1654327 activates OBs 14 and 47. For a binary model, this gives
1237 = 3.0

1654327 = 2.0.

Under a graded model, the difference between the primes is even larger, as OBs 14 and 47 are activated to a low level and have low connection weights. For example, under the above parameters for $B1$ and $B2$,

$$1237 = 2 \times (1.0 \times 1.0) + (0.8 \times 0.8) = 2.64,$$

while

$$1654327 = 2 \times (0.4 \times 0.4) = 0.32.$$

Hence any OB scheme predicts stronger facilitation (faster reaction times, RTs) for 1237 than 1654327.

Under the spatial-coding model,

$$1654327 = 3 + 2 \times S2$$

because letters 1, 4, and 7 match the target exactly, while letters 3 and 5 are shifted by two positions, and

$$1237 = 3$$

because letters 1, 2, and 3 match exactly. Thus the spatial-coding model yields a larger lexical input for 1654327 than 1237, and therefore predicts greater priming effects for the former condition than the latter.

Thus the OB and spatial-coding models give different predictions for 1237 versus 1654327, which we denote as subset and reversed interior primes, respectively. The OB scheme predicts an advantage for subset primes, while the spatial-coding scheme predicts an advantage for reversed interior primes. In the following experiment, we directly compare RTs for these two types of primes as a preplanned comparison over and above the standard analysis of main effects and interactions.

To minimize phonological effects, we use a 33 ms unmasked prime. Grainger et al. (2006) showed that such primes provide robust orthographic, but not phonological, facilitation. In order to further control for any possible phonological or positional advantage for matching the first three contiguous letters in the prime 1237, we also include 1567 in the subset condition. Predictions are the same for 1237 and 1567.

However, the influence of a prime is not limited to excitatory effects. We must also consider sources of inhibition, which could emerge from bottom-up input, or lateral inhibition from nontarget words. We next consider whether any observed advantage for a subset prime over a reversed interior prime could be a result of greater inhibition by the reversed interior prime (rather than increased excitation by the subset prime), although it should be noted that inhibitory priming effects are typically not observed with such a brief prime duration (Grainger, 1992).

An implemented version of the spatial-coding model (Davis, 2010) includes two sources of bottom-up inhibitory input to a word node: inhibition if the stimulus length mismatches the length of the word (not dependent on the degree of difference in length), and inhibition from each mismatching letter. A letter is considered mismatching if it is not in the word, or if it is present but shifted by more than two positions. For example, for the word JUDGE, the inputs jud and judg each yield a length mismatch and no letter mismatches, while eudgi and audgk each yield two-letter mismatches and no length mismatch. The inhibitory costs of length and letter mismatches are necessarily both small (and therefore roughly equivalent), in order to allow strong activation of the target by insertion, deletion, and replacement neighbors. Note that the subset prime yields a length mismatch and one letter shifted by more than two positions, while the reversed interior prime yields two letters shifted by more than two positions. Hence subset and reversed interior primes yield the same number of bottom-up inhibitory sources (i.e., two in each case).

The amount of lateral lexical inhibition depends on the degree to which nontarget words are activated by the prime, as well how strongly the target reactivates these competitors. For example, a pseudoword neighbor prime that shares no neighbors with the target is more facilitative than one that shares neighbors, because the target strongly reactivates shared neighbors, leading to more inhibition of the target (van Heuven, Dijkstra, Grainger, & Schriefers, 2001). Hence, to control for lateral inhibition, it is necessary to control both for the total lexical activity created by the prime, and the degree to which the target reactivates these lexical items.

To address the issue of possible increased total lexical activation for reversed interior primes (due to more diffuse lexical activity in response to a longer prime), we include control primes 1dd7 and 1ddddd7. Equivalent RTs for these two primes would indicate that any advantage for subset primes is not simply a result of greater lexical inhibition from the longer reversed interior primes. Equivalent RTs for these primes would also eliminate the possibility that any advantage for the subset prime is simply due to greater visibility of the external letters or less masking of the target with a shorter prime.

As for controlling for the degree to which the target reactivates the primed competitors, both models predict stronger reactivation for the subset prime. A subset prime consists of three unshifted letters and one letter shifted by three positions, yielding three OBs in the target and three OBs that are not. A reversed interior prime consists of three unshifted letters, two letters shifted by two positions, and two letters shifted by four positions, creating two OBs that are in the target and nine OBs that are not. That is, reversed interior primes contain a higher ratio of shifted to unshifted letters (or non-matching to matching OBs) than subset primes. Therefore, a reversed interior prime is more likely to activate words based on shifted letters/mismatching OBs; such words will be less strongly activated by the target than words primarily containing matching letters/OBs. Hence, if anything, reversed interior primes should provide less sustained inhibition than subset primes.

Therefore, to control for possible differences in total initial lexical activation, exterior letter visibility, and target masking under primes of different lengths, we include a comparison between 1dd7 and 1ddddd7. For other inhibitory sources,
the subset prime is expected to yield equivalent or increased inhibition with respect to the reversed interior prime. Hence, if we observe equivalent RTs for 1dd7 and 1dddd7, but faster RTs for subset than reversed interior primes, the advantage for subset primes is due to greater bottom-up input.

Summing up the predictions to be tested in the present study, the letter position coding mechanisms used in open-bigram coding and the SOLAR model make opposing predictions with respect to two specific priming conditions: subset primes and reversed interior primes. This allows us to implement a strong inference strategy in the manner of Platt (1964), where rather than testing the prediction of a single model or hypothesis, we can test the differing predictions derived from two models. Most critical is the fact that OB models predict stronger priming from subset primes than reversed interior primes, whereas the SOLAR model predicts stronger priming from reversed interior primes than subset primes.

Methods

Participants

Thirty-two students of the University of Provence participated in the experiment for course credit. All were native speakers of French and reported having normal or corrected-to-normal vision.

Design and Stimuli

One hundred twenty French words of 7 letters in length served as targets (see Appendix for the complete list of word targets and corresponding prime stimuli: http://gsite.univ-provence.fr/gsite/Local/lpc/dir/grainger/Whitney-Bertrand-Grainger-Appendix.pdf). The word targets had an average frequency of 38 occurrences per million (range 1–380; New, Pallier, Brysbaert, & Ferrand, 2004), and an average orthographic neighborhood size of 2.03 (range 1.25–3.15) measured with the OLD20 metric (Yarkoni, Balota, & Yap, 2008). An equivalent number of 7-letter nonword targets were included for the purposes of the lexical decision task. These were all orthographically legal and pronounceable letter strings in French. Targets were selected with the aim to minimize within-word letter repetitions in general, and particularly at positions 1237 or positions 1567, which were used to form subset primes. Each target stimulus was associated with four different prime stimuli: (1) subset primes formed by removing three inner letters from the target and concatenating the remaining letters which were the first and last letters plus either the second and third letters or the fifth and sixth letters (1237, 1567); (2) the control primes for the subset condition formed by the first and last letters of the target and two letters not in the target (1dd7); (3) reversed interior primes formed by the first and last letters of the target in the correct position and the inner letters in reversed order (1654327); and (4) the control primes for the reversed order condition (1dddd7). Thus, the factor Prime Type (subset vs. reversed interior) was crossed with Relatedness (related vs. control) in a 2 × 2 factorial design. Four stimulus lists were created each containing all target stimuli equally divided into the four priming conditions, and with targets associated with different primes across the four lists. Participants saw only one list, and an equal number of participants were assigned to each list in a replicated Latin-square design.

Procedure

Each trial was composed of the following series of events. Two fixation bars presented above and below the center of the screen were presented for 500 ms followed 500 ms later by the prime stimulus for 33 ms, which was immediately replaced by the target until participants responded or a timeout of 4,000 ms. Primes and targets were presented in white lowercase letters on a black background at a viewing distance of approximately 60 cm. Primes were in 10-point and targets in 12-point Arial font. Participants were instructed to fixate between the fixation bars and to indicate if the string of letters that remained on the screen was a word in French by pressing one button on a response box with the index finger of their preferred hand, and another button with the other hand if the stimulus was not a word. Participants were instructed to respond as rapidly and as accurately as possible.

Results

Correct RTs to word targets were analyzed following removal of outliers (RTs > 1,500 ms) using a Linear Mixed Effect analysis (“Proc Mixed” procedure of SAS Version 9.1) treating participants and items as random factors, and Relatedness and Type of Prime as fixed factors (restricted maximum likelihood estimation was used to estimate the parameters). Estimated mean RTs and percentage errors to word and nonword targets are shown in Table 1.

In the analysis of correct RTs to word targets there was a significant main effect of Relatedness, F(1, 1902) = 4.19; p = .04, a main effect of Type of Prime, F(1, 1902) = 18.90; p < .0001, and a significant interaction, F(1, 1902) = 5.35; p = .02. Follow-up analyses showed that RTs were significantly shorter following related subset primes than the corresponding control primes, t(1902) = 3.10; p = .002, and the related subset primes generated faster RTs than the related reversed interior primes, t(1902) = 4.69; p < .0001. There was no significant priming effect for the reversed interior primes relative to the respective control primes, t < |1|, and no significant difference between the two unrelated conditions, t(1902) = −1.45; p = .14.

For the analysis of percentage errors to word targets we applied the “proc glimmix” procedure for binary distributions (SAS, Version 9.2). This revealed a significant main effect of Relatedness F(1, 2157) = 4.53; p = .0335, a marginally significant effect of Type of Prime, F(1, 2157) = 3.33; p = .068, and no interaction, F(1, 2157) = 2.31; p = .1287. Finally, there was no significant difference (6 ms in RT and 1.5% difference in error rate) between
the two related subset prime conditions (1237 vs. 1567, \(ts < |1|\)).

An analysis of correct RTs to nonword targets revealed a main effect of Relatedness, \(F(1, 1913) = 3.78; p = .05\), no main effect of Type of Prime, \(F(1, 1913) < 1\), but a significant Relatedness \(\times\) Type of Prime interaction, \(F(1, 1913) = 7.29; p = .007\). Follow-up analyses revealed a significant priming effect for the reversed interior primes relative to the seven-letter control primes, \(t(1913) = 3.30; p = .001\), and reversed interior primes relative to related subset primes, \(t(1913) = 2.43; p = .015\), but no priming effect for the subset primes relative to the four-letter control primes, \(t < |1|\). There was only a marginally significant effect of Type of Prime, \(F(1, 2157) = 3.29; p = .0697\), in an analysis of the percentage errors to nonword targets.

### Discussion

Word targets following subset primes were responded to 26 ms more rapidly than word targets following reversed interior primes, while RTs did not differ significantly between the two control conditions (1d7d and 1dddd7). Furthermore, there was a significant 17 ms priming effect for subset primes (1237/1567) relative to the corresponding control primes and a nonsignificant 2 ms effect for reversed interior primes (1654327), and the critical interaction between Relatedness and Type of Prime was significant in the analysis of RTs to word targets. This pattern of results is clearly in favor of the predictions of OB models and against the predictions of spatial coding. The OB coding scheme correctly predicts this pattern because it explicitly encodes letter order.

The pattern of results seen in RTs to nonword targets showed the opposite pattern, with significant 28 ms priming effect for reversed interior primes, and a nonsignificant 7 ms inhibitory effect for subset primes. This result suggests that the reversed interior primes enhanced the evidence that the target was a nonword, possibly by lowering the evidence that the target is a word (e.g., Dufau, Grainger, & Ziegler, 2011). Most important, this result suggests that the reversed interior primes were not creating more lexical activation via shared neighbors (see Introduction section), since this should have generated more evidence for a word, which would have produced an inhibitory priming effect on nonword targets. In any case, this particular result certainly merits further exploration, as is the case for masked orthographic priming effects on nonword targets in general.

The current results are based on inherent predictions of the two models under investigation, which are independent of parameter values. This was not the case in previous attempts to experimentally differentiate between the models (Davis & Bowers, 2006), in which the authors concluded that their results provided evidence against OB models. However, as discussed in Whitney (2008), the authors’ conclusion was specific to binary OBs, and the initial parameterization of the SERIOL model in which OB activations decreased with increasing string position (Whitney, 2001). Their experiments do not distinguish between the spatial-coding model and graded OBs in which activation depends only on letter separation, as is the case in the current parameterization of the SERIOL model (Whitney, 2004), and in other graded OB models. In contrast, the present experiments are based on predictions that do not depend on particular parameterizations of the models under consideration.

Another reason to prefer models that implement OB input coding is that the theories differ in the plausibility of their lexical input functions. In neural modeling, a connection weight represents the efficacy of a synapse (or a set of synapses). The larger the connection weight, the better the transfer of the signal from the sending neuron(s) to the receiving neuron(s). This is modeled by multiplying the activation of the sending unit by the connection weight, to obtain the input to the receiving unit. In models that use an OB scheme, lexical input is based on these standard assumptions. Furthermore, we consider OBs to be a reading-specific adaptation of a known mechanism of visual encoding, wherein neurons are tuned to combinations of object parts, which are not necessarily contiguous (Brincat & Connor, 2004).

However, the spatial-coding input function relies on direct comparisons between letters’ positions in the input and the target. Davis (2010) specifies that this comparison is carried via a temporal encoding of letter order with delay lines to position-specific comparator banks. Each word requires a bank of letter processors for each letter position in the word, and each bank of processors involves multiple letter units corresponding to the same letter coming from retinal locations. This specification is quite complex, and it is unclear that the sketched neural mechanism can actually

### Table 1. Summary of experimental results. Estimated mean and standard error (SE) or RTs and percentage errors to word and nonword targets following related subset (4-letter) and reversed interior (7-letter) primes and the corresponding unrelated primes.

<table>
<thead>
<tr>
<th></th>
<th>Subset (4 letters)</th>
<th>Reversed interior (7 letters)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean RT</td>
<td>SE</td>
</tr>
<tr>
<td>Words</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Related</td>
<td>601</td>
<td>15.8</td>
</tr>
<tr>
<td>Unrelated</td>
<td>618</td>
<td>15.8</td>
</tr>
<tr>
<td>Nonwords</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Related</td>
<td>710</td>
<td>21.2</td>
</tr>
<tr>
<td>Unrelated</td>
<td>703</td>
<td>21.2</td>
</tr>
</tbody>
</table>
implement the proposed input function. In contrast, no word-specific letter units are necessary in an open-bigram encoding, since open-bigrams directly connect to whole-word representations, and the neural implementation of the proposed input function is straightforward. Hence, the spatial-coding scheme is computationally much more costly than open-bigrams, while offering no advantage in explanatory capacity.

Acknowledgment

This research was supported by ERC advanced Grant No. 230313 awarded to J. Grainger.

References


Received October 7, 2010
Revision received July 25, 2011
Accepted July 27, 2011
Published online November 4, 2011

Jonathan Grainger
LPC – Pôle 3C
Université de Provence
3 place Victor Hugo
13331 Marseille France
Tel. +33 (0)4 13 55 09 99
E-mail jonathan.grainger@univ-provence.fr