In the present study, online measures of letter identification were used to test computational models of letter perception. Event-related potentials (ERPs) were recorded to letters and pseudoletters revealing a transition from feature analysis to letter identification in the 100–200-ms time window. Measures indexing this transition were then computed at the level of individual letters. Simulations with several versions of an interactive-activation model of letter perception were fitted with these item-level ERP measures. The results are in favour of a model of letter perception with feedforward excitatory connections from the feature to the letter levels, lateral inhibition at the letter level, and excitatory feedback from the letter to the feature levels.

Keywords: Letter perception; Event-related potential; Computational modelling; Feedback; Lateral inhibition.

In our literate societies, letters are simple, easy-to-control, highly functional, and overlearned two-dimensional visual patterns associated with a single name. They also gather a dose of complexity due to their supposed componentiality, their multiple form instantiations, and their similarities. Letters therefore offer a paradigm case for the investigation of visual pattern recognition in humans, incorporating most of the complexities while remaining experimentally and theoretically tractable (for a recent review, see Grainger, Rey, & Dufau, in press).

Paradoxically, although computational modelling of visual word recognition is one of the most active domains of cognitive psychology, little attention has been devoted to computational modelling of letter perception (see, however, Larsen & Bundesen, 1996; Shevtsova & Reggia, 1999, 2000, 2002). Indeed, most models of word recognition only start to examine the processes involved once letter perception is complete (e.g., Ans, Carbonnel, & Valdois, 1998; Harm & Seidenberg, 1999, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996; see, however, Ziegler, Rey, & Jacobs, 1998). The interactive activation model of McClelland and Rumelhart (1981) and its derivatives (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Grainger &
Jacobs, 1996; Jacobs & Grainger, 1992; Jacobs, Rey, Ziegler, & Grainger, 1998; Perry, Ziegler, & Zorzi, 2007) include feature and letter representations, but processing at these levels is generally ignored. Notably, in these models, it is assumed that (a) both excitatory and strong inhibitory connections link the feature to the letter level, (b) no feedback is sent from the letter level to the feature level, and (c) although lateral inhibition between letter representations is posited in the original interactive activation model, the lateral inhibition parameter is always set to zero, indicating finally that no letter-to-letter inhibition is assumed at this level. One goal of the present study was therefore to address these undefended assumptions by testing different computational models of letter perception against online measures of letter processing.

Errors obtained in various perceptual identification tasks have been one major dependent measure in studies of letter perception, often used to generate confusion matrices (for an extensive review of this literature, see Mueller, Weidemann, & Shiffrin, 2008). Standard behavioural measures of response times have also been recorded and have frequently been combined with the priming technique (Arguin & Bub, 1995; Bowers, Vigliocco, & Haan, 1998; Courrieu, Farioli, & Grainger, 2004; Grainger & Jacobs, 1991; Jacobs & Grainger, 1991; Petit & Grainger, 2002; Ziegler, Ferrand, Jacobs, Rey, & Grainger, 2000). More recently, these behavioural observations have been complemented by the direct recording of brain activity using magnetoencephalography (MEG; Tarkiainen, Cornelissen, & Salmelin, 2002; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999) and event-related potentials (ERPs; Petit, Midgley, Holcomb, & Grainger, 2006).

The overall pattern of effects observed in the MEG and ERP studies revealed an occipital activation at 100 ms after stimulus onset that was not sensitive to the specific content of the stimulus and that was interpreted as reflecting low-level visual feature processing. Subsequent inferior occipito-temporal activation was found at around 150 ms poststimulus onset and was interpreted as reflecting the earliest stage of stimulus-specific processing. This hierarchical interpretation of the results is consistent with the hierarchical organization of the visual system (e.g., Felleman & Van Essen, 1991; Rolls, 2000, 2007; Van Essen, 2005) and the presence of object- and category-selective neurons in higher visual areas (Kreiman, Koch, & Fried, 2000; Logothetis, Pauls, & Poggio, 1995; Perrett, Rolls, & Caan, 1982; Quiroga, Reddy, Kreiman, Koch, & Fried, 2005). The 150-ms time window is also consistent with several ERP studies showing that a similar amount of time is needed to begin the identification of a visual target (e.g., an animal or a vehicle) in a natural scene (e.g., Schendan, Ganis, & Kutas, 1998; Thorpe, Fize, & Marlot, 1996; for a review, see VanRullen, 2007). In their masked prime ERP study, Petit et al. (2006) gave slightly later estimates of the timing of letter perception processes (starting around 180 ms posttarget onset), but this is likely due to general interference generated by the mask and prime stimuli in this particular paradigm.

Overall, these data suggest that letter identification progressively takes place within a transitory 100–200-ms time window, and one can assume that the dynamics of this transition varies across individual letters. By looking at the ERP properties of this time window for individual letters, it should therefore be possible to obtain an online measure of item-specific letter identification in the brain. The first—empirical—goal of the present study is precisely to record ERPs for a restricted set of letters that will be sufficiently repeated to extract from the ERP signal a stable measure differentiating each letter throughout the 100–200-ms time window.

The second—theoretical—goal is to use these variations in ERPs to individual letters in order to test the predictions of different computational models of letter perception. The core representational assumptions of these models are derived from the interactive-activation model of McClelland and Rumelhart (1981) and are represented in Figure 1. A hierarchical organization is assumed with two levels of processing: a
Localist representations of features and letters are used at each processing level as simplified instantiations of the pattern of activity related to elementary, feature-based visual processes on the one hand and higher level, item-specific identification processes on the other (Page, 2000). These representations are interconnected by feedforward, feedback, and lateral connections, each being characterized by a fixed parameter that determines its weight.

There are feedforward excitatory and inhibitory connections from the feature to the letter level, and parameters $\alpha_{FL}$ and $\gamma_{FL}$ determine their respective weight. Lateral inhibitory connections at the letter level and feedback excitatory connections from the letter to the feature level are characterized by parameters $\gamma_{LL}$ and $\alpha_{LF}$, respectively. By systematically varying these parameter values (including the null value), we test the predictions of the different computational instantiations of this general architecture against the ERP data. More specifically, we test the adequacy of three processing assumptions: (a) feedforward inhibitory connections, (b) lateral inhibition, and (c) excitatory feedback.

In the next sections, we first present the experiment that allowed us to collect ERP measures at the level of individual letters. By comparing the average ERP waveforms for letters and pseudoletters, we show that these data replicate previous findings obtaining a difference between these categories starting around 150 ms poststimulus onset. Then, we explain how we computed a derived ERP measure that indexes the 100–200-ms transition phase from features to letter identification for each individual letter. This measure is then used to test simulations from the different computational models that were obtained by varying the feedforward inhibitory connection ($\gamma_{FL}$), the lateral inhibitory connection ($\gamma_{LL}$), and the excitatory feedback loop ($\alpha_{LF}$) in the generic interactive-activation model of letter perception.

ERP EXPERIMENT

The purpose of this experiment was twofold. First, following the logic of Thorpe et al. (1996), ERPs for a set of letters and pseudoletters were obtained by averaging a large number of trials per participant, maximizing the signal-to-noise ratio. By matching the sets of letters and pseudoletters at the level of visual features, it was further assumed that low-level visual processing should lead on average to the same ERP signal at the earliest stages of processing. The time when the two signals start to diverge should then be tied to processes that are specific to letter identification. The goal was therefore to replicate previous findings that reported a divergence point between two visual categories (i.e., targets vs. nontargets) as early as 150 ms. To avoid any interpretation of this divergence in terms of decision-making mechanisms (e.g., Johnson & Olshausen, 2005), a simplified experimental procedure was adopted that guaranteed that attention was directed to the visual stimuli but only occasionally required that the participant make a decision and produce a motor response.

Second, because we were interested in directly testing item-level predictions from computational...
models against ERP measures of individual-letter identification, we used the same letter font in the simulations and in the experiment. This font is constructed using a feature matrix adapted from Rumelhart and Siple (1974) and consisted of a set of 16 independent features that allowed us to characterize any of the 26 letters of the alphabet (for example, letter A is composed of eight features from this matrix: 1, 2, 3, 4, 5, 6, 10, and 12—see Figure 2 for a representation of the feature matrix). This choice also allowed us to match letters and nonletters in terms of their component features. To obtain reliable letter level ERPs, we decided to reduce the set of tested letters in order to maximize the number of trials per letter while keeping the experiment duration acceptable for participants. From the 26 letters of the alphabet, we selected 14 letters that were only composed of horizontal and vertical lines in the Rumelhart and Siple (1974) font while excluding letters having a nonstandard shape within this font (such as letters G and J). Restricting the number of target letters therefore allowed us to repeat each letter many times increasing the quality of the resulting ERP signal at the level of individual letters.

Method

Participants
A total of 20 healthy volunteers from the University of Provence took part in the experiment as paid volunteers (average age 22.9 years, 6 women). All participants reported being right-handed native speakers of French with normal or corrected-to-normal vision.

Material
Figure 2 shows the 14 letters and the 14 pseudoletters used in the ERP experiment together with the feature matrix from which these items were constructed. The letters were only composed of vertical and horizontal lines. The pseudoletters were constructed with the two following constraints: (a) The same set of features composing the letters was used to construct the pseudoletters (i.e., 92 features were needed to compose the set of letters, and exactly the same 92 features were

Figure 2. Left side: the feature matrix used to construct the letters and pseudoletters (adapted from Rumelhart & Siple, 1974). Right side: the 14 letters and pseudoletters derived from this matrix.
used for pseudoletters); and (b) the distribution of the number of features across pseudoletters was the same as that for letters (i.e., there were two letters/pseudoletters with four features, one letter/pseudoletter with five features, four letters/pseudoletters with six features, two letters/pseudoletters with seven features, four letters/pseudoletters with eight features, and one letter/pseudoletter with nine features). Thus the two sets of stimuli were on average composed of exactly the same visual features. Stimuli were displayed on a 17” cathode ray tube (CRT) monitor (placed at 90 cm from participants) as white signs of 100 × 100 pixels (=4 cm) on an 800 × 600-pixel black background (32 × 24 cm).

Procedure
After completing informed consent, participants were seated comfortably in a sound-attenuated and dimly illuminated room. The experiment started with two successive presentations of the complete set of letters. To check that participants correctly identified the letters, they had to name each letter (presented one after the other for a duration of 2,000 ms). Instructions concerning the experiment were then given to participants. They were informed that a letter or a pseudoletter would appear at the centre of the screen. Their task was simply to remember the item. Occasionally, after the presentation of an item, a green circle was displayed indicating that they would have to compare this item with an upcoming item (noncritical comparison trials). In this case, participants were expected to press a right button if the two items were identical and a left button otherwise (participants held two small button boxes placed in their hands allowing responses to be given with the thumb).

A regular trial was composed of the following sequence. A fixation point (i.e., an asterisk, “*”) was presented for 500 ms, followed by a blank during 300 ms, the target item (letter or pseudoletter) for 500 ms, and a blank intertrial interval of 500 ms. On comparison trials, this sequence was completed with the following events: a green circle appeared for 500 ms, followed by a blank during 500 ms, the to-be-compared item (letter or pseudoletter) for 500 ms, a blank for 500 ms, and the word “response” during 2,000 ms (participants had to produce their response during this period), and, finally, the next trial started after a blank of 1,000 ms. There were 28 comparison trials (14 with identical items and 14 with different items) that appeared randomly during the experimental session. These trials were not included in the ERP analyses. Each letter and pseudoletter from regular trials appeared randomly 20 times during the experiment leading to a total of 280 letter trials + 280 pseudoletter trials + 28 comparison trials = 588 trials. The experimental session was organized in three blocks (i.e., with two pauses) and lasted approximately 30 minutes.

ERP recording
The electroencephalography (EEG) activity was recorded continuously through the Active Two Biosemi system from 64 electrodes mounted on an elastic cap (Electro-Cap Inc.) that was positioned according to the 10–10 International system (American Clinical Neurophysiology Society, 2006). Two additional electrodes (CMS/DRL near PZ) were used as online reference (for a reference description, see www.biosemi.com; Schutter, Leitner, Kenemans, & van Honk, 2006). The montage included 10 midline sites (FPZ, AFZ, FZ, FCZ, CZ, CPZ, PZ, POZ, OZ, and IZ) and 27 sites over each hemisphere (FP1/FP2, AF3/AF4, AF7/AF8, F1/F2, F3/F4, F5/F6, F7/F8, FC1/FC2, FC3/FC4, FC5/FC6, FT7/FT8, C1/C2, C3/C4, C5/C6, T7/T8, CP1/CP2, CP3/CP4, CP5/CP6, TP7/TP8, P1/P2, P3/P4, P5/P6, P7/P8, P9/P10, PO3/PO4, PO7/PO8, and O1/O2). Four additional electrodes were used to monitor eye movements and blinks (two placed at lateral canthi and two below the eyes), and two additional electrodes were used for an offline re-referencing (placed behind the ears on mastoid bone). Continuous EEG was digitized at 256 Hz and was filtered offline (20 Hz low-pass, 24 dB/octave). The low-pass filter removed high-frequency interference on the latencies and
amplitudes of the peaks of interest, thus enhancing the quality of the ERP measure. Recording obtained from a left mastoid electrode was used offline to re-reference the scalp recordings to the left mastoid.\textsuperscript{2} Data were epoched on target presentation: (−100, 800) ms epoch, (−100, 0) ms baseline. Epochs with eye movements, blinks, or electrical activities greater than ±50 μV were rejected (19.7% of the trials). Two participants were discarded from further analysis due to excessive movement artefacts.

**Behavioural results and ERP analyses**

The percentage of correct responses in the comparison task was on average 94.8% (SD: 4.74; minimum = 86%; maximum = 100%). Because no response was recorded on critical trials, the fact that participants had a high level of performance on noncritical trials is a guarantee that they adequately perceived and identified the target stimuli.

As for the ERPs, we first report the comparison between the averaged ERPs obtained for letters and pseudoletters. Since these two sets of items were matched on low-level visual dimensions, the earliest difference in the ERP is taken to be concomitant with the brain activity related to letter identification.

This comparison was computed by averaging ERPs across all letter targets on the one hand and all pseudoletter targets on the other, across all participants. The difference maps between letters and pseudoletters are displayed in Figure 3a for three different time-windows (90–130, 130–170, and 170–210 ms). No difference can be seen in the first map, but a clear difference appears in the second and third maps at fronto-central electrodes. To maximize the signal-to-noise ratio, the mean ERP across 13 central electrodes was calculated for each participant in the letter and pseudoletter conditions (Figure 3b displays the set of selected electrodes). The corresponding grand average for letters and pseudoletters is presented in Figure 3c. A repeated measure analysis of variance (ANOVA) was used to analyse these central ERPs using type-of-item (letters vs. pseudoletters) as a within-participant independent variable and mean ERP amplitude as the dependent variable. A continuous time bin analysis was employed to determine the onset of the divergence between the letter and pseudoletter ERPs. The sliding time-window width was 27 ms (time-sampled point: ±3 points) and was applied for each time-sampled point in the ERP recording. The onset of the divergence between the letter and pseudoletter ERPs was determined when at least 15 consecutive $F$ values exceeded a $p = .05$ level of significance (Rugg, Doyle, & Wells, 1995). In the present analysis, the onset was reached by 145 ms: 4.88 < $F(1, 17) < 11.12$, for the 15 consecutive time-sampled points.

This global analysis was followed by a second ERP analysis performed at the level of individual letters. Figure 4 shows the mean ERP for each of the 14 letters during the 300 ms that follow stimulus onset (together with the grand mean for pseudoletters as a reference\textsuperscript{3}). Although the average ERP signals for each letter do display some variation, visual inspection reveals a consistent pattern of activity in the 100–200-ms time-window. This pattern is characterized for all letters by a peak of negativity around 100 ms (i.e., $N_{100}$ or $N_1$) followed by a systematic transition and a positive peak around 200 ms (i.e., $P_{200}$ or $P_2$). Given the result of the grand average letter/pseudoletter comparison, one can assume that individual-letter identification takes

\textsuperscript{2} Following Thorpe et al. (1996), we used the mastoid as the reference electrode. Although this choice does affect the spatial distribution of ERPs, it does not affect the timing of the difference between letters and pseudoletters. Since we were interested in getting precise information about timing (and not localization), we decided to use the same reference as that in the study by Thorpe et al. (1996).

\textsuperscript{3} Note that the grand mean for pseudoletters is only used here as a common reference in order to visualize variability between ERPs for individual letters. A direct comparison between this ERP grand mean and individual-letter ERPs is meaningless because the whole set of pseudoletters was not matched on visual features with each of the individual letters.
place, on average, between these N100 and P200 peaks of activity. We therefore decided to extract a representative index of this N100/P200 transition by measuring the latency and amplitude of the N100 and P200 peaks for each letter. We then averaged the latencies for N100 and P200 peaks for each letter. The amplitude value corresponding to the resulting average latency value was then extracted from the average ERP signal for each letter.

The results of this analysis are presented in Table 1. The latency and amplitude measures provide a characterization of the brain activity for each letter in the time-window associated with the transition from feature analysis to letter identification. Individual-letter latencies and amplitudes were, on average, 111.9 ms and −2.5 mV for N100, and 191.4 ms and 5.7 mV for P200. The critical mean latency and amplitude values were, on average, 151.6 ms and 2.44 mV. These mean values were transformed into z-scores and were used as dependent variables to compare with latency and activation values derived from computer simulations run on different versions of the generic model of letter perception.

**COMPUTATIONAL MODELLING**

As depicted in Figure 1, the present modelling approach is based on the generic framework of the interactive-activation (IA) model proposed by McClelland and Rumelhart (1981). Of course, this framework is an oversimplification of the brain structures and mechanisms involved in
letter perception but it is intended to capture their main properties. The IA framework indeed provides at least a plausible functional simplification of this highly complex biological system, which can allow us to explore some of its dynamical properties (such as connectivity assumptions).

The IA framework uses nodes that are activated when some patterns or properties of the external world are visually presented. In that sense, these nodes are representations, a representation being defined as a pattern of neural activity that is produced by the perceptual system when a given stimulus (e.g., a letter) is visually processed. Even if this pattern of neural activity is necessarily distributed among several neuronal populations, a localist node can nevertheless be used as a functional tool to represent this pattern of activity at a given level of processing. Although the hierarchical organization of these nodes (i.e., with a feature and a letter level of processing) is also a simplification, it is nevertheless consistent with the hierarchical organization proposed by recent models of visual object perception in computational neuroscience (e.g., Rolls, 2007; Van Essen, 2005).

Within the IA framework, the presentation of a given letter stimulus activates feature nodes, which in turn send activation to letter nodes. Time is segmented into processing cycles, and activation of letter nodes increases nonlinearly up to an asymptotic value. A letter is said to be identified when its corresponding node has reached a certain amount of activation (e.g., Grainger & Jacobs, 1996). Letter identification can therefore be characterized by an activation value reached after a certain level of processing.

Figure 4. Event-related potentials (ERPs) for each of the 14 target letters used in the experiment. The dashed line corresponds to the grand average ERP for pseudoletters. These means correspond to the averaged ERPs over the 13 electrodes displayed in Figure 3.
In the present simulations, we test the activation and processing time values generated by different versions of the IA model for the 14 letters used in the experiment with the respective individual-letter mean amplitude and latency ERP values reported in Table 1. As for the ERP values, activation and cycle values obtained for each letter were transformed into z-scores for each version of the IA model.

The different versions of the IA model were obtained by varying systematically three connectivity parameters: feature-to-letter inhibition, letter-to-letter inhibition, and letter-to-feature excitatory feedback, everything else being maintained constant. From these versions, four categories of models were considered: (A) a fully connected model with various weight values for the feature-to-letter inhibition connection, the letter-to-letter inhibition connection, and the letter-to-feature excitatory feedback connection (all these parameter values being different from zero); (B) a model without feature-to-letter inhibition (the parameter for this connection weight being set at zero, i.e., $\gamma_{FL} = 0$); (C) a model without letter-to-letter inhibition (i.e., $\gamma_{LL} = 0$); (D) a model without feedback (i.e., $\alpha_{FL} = 0$). For models B, C, and D, while one of the three considered connections was set at zero, the other connections were systematically varied as for Model A. The descriptive adequacy of these four model categories was evaluated by examining the root mean square deviations between the letter identification values (i.e., activation and processing cycle) generated by these models for the 14 experimental letters and the mean ERP data (i.e., amplitude and latency) from Table 1.

**Method**

The present generic IA model of letter perception exactly follows the general computational principles of the original IA framework, although some adaptations were made in order to implement lateral inhibition at the letter level and feedback from the letter to the feature level. The feature and letter levels are composed, respectively, of 16 feature nodes (see Figure 2) and 26 letter nodes (even if only 14 letters were actually tested in this study). Each node is characterized by an activation value that is updated at every processing cycle.

### Table 1. Latencies and amplitudes for the N100, the P200, and the resulting mean values from the ERPs of individual letters

<table>
<thead>
<tr>
<th>Letter</th>
<th>$N_{100}$ Latency</th>
<th>$N_{100}$ Amplitude</th>
<th>$P_{200}$ Latency</th>
<th>$P_{200}$ Amplitude</th>
<th>Mean Latency</th>
<th>Mean Amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>109</td>
<td>-2.42</td>
<td>191</td>
<td>6.48</td>
<td>150</td>
<td>2.49</td>
</tr>
<tr>
<td>B</td>
<td>105</td>
<td>-1.26</td>
<td>191</td>
<td>8.29</td>
<td>148</td>
<td>4.73</td>
</tr>
<tr>
<td>C</td>
<td>109</td>
<td>-3.62</td>
<td>188</td>
<td>5.75</td>
<td>148.5</td>
<td>1.18</td>
</tr>
<tr>
<td>D</td>
<td>109</td>
<td>-1.58</td>
<td>191</td>
<td>7.87</td>
<td>150</td>
<td>3.98</td>
</tr>
<tr>
<td>E</td>
<td>113</td>
<td>-1.38</td>
<td>199</td>
<td>5.67</td>
<td>156</td>
<td>4.20</td>
</tr>
<tr>
<td>F</td>
<td>113</td>
<td>-2.19</td>
<td>184</td>
<td>6.39</td>
<td>148.5</td>
<td>1.74</td>
</tr>
<tr>
<td>G</td>
<td>113</td>
<td>-2.58</td>
<td>168</td>
<td>4.28</td>
<td>140.5</td>
<td>1.40</td>
</tr>
<tr>
<td>I</td>
<td>113</td>
<td>-1.61</td>
<td>191</td>
<td>5.00</td>
<td>152</td>
<td>3.77</td>
</tr>
<tr>
<td>L</td>
<td>125</td>
<td>-4.62</td>
<td>219</td>
<td>3.98</td>
<td>172</td>
<td>1.37</td>
</tr>
<tr>
<td>O</td>
<td>109</td>
<td>-1.67</td>
<td>199</td>
<td>5.97</td>
<td>154</td>
<td>0.81</td>
</tr>
<tr>
<td>P</td>
<td>117</td>
<td>-3.32</td>
<td>184</td>
<td>4.67</td>
<td>150.5</td>
<td>2.09</td>
</tr>
<tr>
<td>S</td>
<td>109</td>
<td>-3.37</td>
<td>180</td>
<td>4.58</td>
<td>144.5</td>
<td>1.64</td>
</tr>
<tr>
<td>T</td>
<td>105</td>
<td>-2.28</td>
<td>199</td>
<td>4.68</td>
<td>152</td>
<td>1.65</td>
</tr>
<tr>
<td>U</td>
<td>117</td>
<td>-2.59</td>
<td>195</td>
<td>6.34</td>
<td>156</td>
<td>3.10</td>
</tr>
<tr>
<td>Mean</td>
<td>119.9</td>
<td>-2.5</td>
<td>191.4</td>
<td>5.7</td>
<td>151.6</td>
<td>1.6</td>
</tr>
</tbody>
</table>

*Note:* ERP = event-related potential. $N_{100}$ denotes peak of negativity around 100 ms. $P_{200}$ denotes positive peak around 200 ms.
A simulation was initiated by presenting a letter to the model and by setting to one the activation value of the feature nodes that compose this letter. For example, when the letter A was presented to the model, Features 1, 2, 3, 4, 5, 6, 10, and 12 received an activation of 1 at Processing Cycle 1. Activation then spread to the letter level via excitatory and inhibitory connections. The weight of the feature-letter excitatory connection was maintained constant, and its parameter value was the same as that in the original IA model (i.e., $\alpha_{FL} = .005$). Equations used to compute the activation of feature and letter nodes are described in the Appendix.

As mentioned above, four versions of the IA model were tested. Model A is a fully connected network, and parameters $\gamma_{FL}$, $\gamma_{LL}$, and $\alpha_{LF}$ (determining the weight of feature-to-letter inhibitory connections, letter-to-letter inhibitory connections, and letter-to-feature excitatory connections, respectively) were systematically manipulated along the following values: .005, .01, .015, ... up to .1. This cross-variation of these parameters led to $20^3 = 8,000$ different versions of this fully connected model. For Model B, there was no feature-to-letter inhibitory connection (i.e., $\gamma_{FL} = 0$) while $\gamma_{LL}$ and $\alpha_{LF}$ varied along the values: .005, .01, .015, ... up to .1. This led to $20^2 = 400$ versions of the “no feature-to-letter inhibition” model. Model C was the “no letter-to-letter inhibition” model ($\gamma_{LL} = 0$), $\gamma_{FL}$ and $\alpha_{LF}$ varying along the values: .005, .01, .015, ... up to .1 (leading also to $20^2 = 400$ versions of this model). Finally, Model D was the “no feedback” model ($\alpha_{LF} = 0$), $\gamma_{FL}$ and $\gamma_{LL}$ varying along the values: .005, .01, .015, ... up to .1. This led again to $20^2 = 400$ versions of this model.

For all these $8,000 + 400 + 400 + 400 = 9,200$ versions of the IA model, activation curves for each of the 14 target letters were computed. Following Jacobs and Grainger (1991), letter identification was said to occur when activation of a given letter node reached 80% of the activation curve asymptote for that letter (activation of the target letter node at Cycle 50 served to determine this asymptotic value). This read-out allowed us to obtain, for each letter, a single activation and cycle value characterizing letter identification. Figure 5 provides two examples of this procedure. These simulations correspond to the activation curves obtain at the letter level for a model with the following parameter set: $\gamma_{FL} = .01$, $\gamma_{LL} = .05$, and $\alpha_{LF} = .05$. The left and right panels are the curves obtained when letters A and S, respectively, were presented to the model.

This procedure generated activation and cycle values for each of the 14 target letters and for each of the 9,200 versions of the IA model. The root mean square deviation (RMSD) between

Figure 5. Activation curves obtained with the IA model when the letter A (left panel) and the letter S (right panel) were presented.
the activation values generated by each model and the amplitude values of the ERPs, on the one hand, and between the cycle values of the model and the latency values of the ERPs, on the other hand, were computed for each version of the model. These RMSDs were used to evaluate the descriptive adequacy of the 9,200 versions of the generic IA model of letter perception.

Simulation results

Figure 6 displays the results of these simulations for the four categories of IA models. Each point in the graphs corresponds to a version of the IA model (i.e., to a given parameter set) for which performance is characterized by two RMSD values: one comparing letter-level model activations and letter-level ERP amplitudes, and the other comparing letter-level cycles and letter-level ERP latencies. The dashed line represents the smallest RMSD values obtained over all simulations on the latency and amplitude dimensions.

The first result was to obtain the lowest RMSDs both on the latency and the amplitude measures for models of Category B—that is, with no feature-to-letter inhibition. The best parameter set for Category B models corresponded to a model having little lateral inhibition (i.e., \( \gamma_{LL} = 0.01 \)) and strong excitatory feedback (i.e., \( \alpha_{LF} = 0.1 \)). For this parameter set, significant correlations were obtained both between the predicted letter activations and between the ERP letter amplitudes, \( r(13) = 0.55, p = 0.039 \), and between the predicted letter cycles and the ERP letter latencies, \( r(13) = 0.52, p = 0.058 \). Consistent with this finding, models having no feedback (i.e., models of Category D) led to the worst RMSDs, suggesting that feedback from the letter to the feature level is an important factor for the dynamics of the present interactive activation model. Similarly, Category C models (i.e., with no lateral inhibition) outperformed Category D models, and this result again suggests that lateral inhibition is probably an important factor but to a lesser extent than the letter-to-feature excitatory feedback. Finally, the best fits were obtained with the fully connected model (i.e., Category A models) for the following parameter set: \( \gamma_{FL} = 0.02, \gamma_{LL} = 0.015, \) and \( \alpha_{LF} = 0.1 \). This parameter set led to marginally significant correlations on latency, \( r(13) = 0.47, p = 0.09 \), and amplitude values, \( r(13) = 0.52, p = 0.06 \). Note that exploring the parameter space up to 0.5 for the three manipulated connection weights did not change this general pattern of results.

DISCUSSION

In the empirical section of the present study, we reported an ERP experiment that revealed two main results. First, by comparing ERPs for a set of letters and pseudoletters that were matched on low-level visual dimensions, we found a divergence point in the ERP signal that appeared 145 ms after stimulus onset. This result is consistent with previous ERP and MEG findings indicating that high-level category-selective visual representations start to be activated at around 150 ms (Large, Kiss, & McMullen, 2004; Schendan et al., 1998; Tarkiainen et al., 2002; Tarkiainen et al., 1999; Thorpe et al., 1996) or 170 ms (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Jeffreys, 1996; Liu, Harris, & Kanwisher, 2002; Low et al., 2003) after stimulus onset. This evidence suggests that processing in the first 100 ms is likely to be dedicated to low-level (feature-based) visual analysis and that letter identification takes place in the 100–200-ms time-window.\(^4\)

Second, we were able to characterize this 100–200-ms transition time-window by computing an ERP measure derived from N\(_{100}/P_{200}\) peaks.

\(^4\) As already noted in the introduction, ERP measures obtained in the masked priming paradigm with letter stimuli led to slightly later estimates (Petit et al., 2006). This is probably due to a general slowing of the target letter identification process following a pattern mask and prime stimulus. This possibility is consistent with the findings of Ziegler et al. (2000) who demonstrated, using the incremental priming technique proposed by Jacobs, Grainger, and Ferrand (1995), that most of the repetition priming effect found for letter targets in an alphabetic decision task was in fact driven by a reduction in the interference effect of unrelated primes.
This integrated, item-specific, ERP measure was then used as an online index of letter identification processes. The use of this measure was motivated by the overall comparison of ERPs to letters versus pseudoletters suggesting that letter-specific processing was already engaged on average at 145 ms—that is, between the N\textsubscript{100} and P\textsubscript{200} peaks. If N\textsubscript{100} is likely to be related to low-level visual processing and P\textsubscript{200} to later processes, such as access to letter names (Petit et al., 2006), then the proposed integrated measure is likely to index the phase of letter-specific processing since its latency values fall, on average, near the divergence point obtained in the comparison between letters and pseudoletters (i.e., 151.6 ms, see Table 1). We would hypothesize that this earliest phase of letter-specific processing involves the mapping of visual features onto letter representations, as described in the generic IA model of letter identification.
The most interesting aspect of this online measure was the possibility of obtaining letter-level variance and to use this variance to test computational models of letter perception. Note that if this measure is tied to the visual identification of letters, the resulting variance will therefore depend on the properties of the letter font adopted in the experiment. That is the reason why it was important to keep the font constant both in the experiment and in the simulations. Within the framework sketched by Petit et al. (2006), the present analysis concerns the first stage of letter perception involving case-specific letter representations. The results of Petit et al. (2006) suggest that this initial phase is rapidly followed by the activation of more abstract, case-invariant letter representations that provide access to higher level phonological and conceptual codes.

The conclusions that can be drawn from the modelling section are twofold. First, for some versions of the IA model of letter perception, there were significant correlations between the ERP measures and individual-letter identification measures derived from the simulations. This result was, a priori, not obvious because of the integrated nature of the ERP signal. Finding such correlations suggests that computational modelling might be a useful tool for interpreting ERP waveforms and for linking this electrophysiological measure to cognitive mechanisms (for a similar approach, see Braun et al., 2006). These significant correlations also provide a validation of the variance resulting from the ERP measures, suggesting that these measures indeed capture some part of the variance related to letter identification processes.

Second, a systematic comparison of the descriptive adequacy of four categories of models showed an advantage of a model composed of feature-to-letter excitatory connections, lateral inhibition, and letter-to-feature feedback, but no feature-to-letter inhibition. As already mentioned, this result is globally consistent with recent propositions in the computational neuroscience of pattern recognition, suggesting that the visual system is organized hierarchically with excitatory feedforward connections, lateral inhibition, and feedback loops (e.g., Bullier, 2001; Dehaene, Cohen, Sigman, & Vinckier, 2005; Rolls, 2000, 2007; Van Essen, 2005). Conversely, it is totally inconsistent with the current description of computational models of word reading derived from the original IA model that assume strong feedforward inhibitory connections, no lateral inhibition at the letter level, and no feedback from the letter to the feature level (e.g., Coltheart et al., 2001; Grainger & Jacobs, 1996; Jacobs & Grainger, 1992; Jacobs et al., 1998; Perry et al., 2007).

Although it could be argued that letter recognition processes are underspecified in models of word recognition because they only account for a small amount of the word recognition variance or because adult word-reading processes are functionally independent from letter recognition processes (Rey & Schiller, 2006), the present results could still be used to elaborate general connectivity principles governing the relations between and within levels of processing. The results are indeed in favour of computational architectures having interlevel feedforward and feedback excitatory connections, together with intralevel lateral inhibitory connections. These connectivity principles would therefore be inconsistent with those adopted by models that allow interlevel connection weights to take negative values through backpropagation learning mechanisms (e.g., Harm & Seidenberg, 2004; Plaut et al., 1996).

In conclusion, the present study used item-level measures from ERPs in order to obtain an online index of individual-letter perception processes. Because letters are highly frequent visual patterns, they can be intensively repeated during an experiment, yielding a stable item-level ERP signal that closely characterizes individual-letter processing. Increasing the set of letters and comparing upper- and lower-case letters and different letter fonts having different visual properties will certainly provide empirical constrains for further tests of computational models of letter perception. The present study also suggests that future directions in computational modelling can certainly benefit from the neurophysiological information available from ERP studies together with studies of the structure and dynamics of the primate visual cortex.
REFERENCES


Equations used for computing feature and letter node activations

At each processing cycle \( t > 1 \), activation of letter node \( L_i \) is equal to:

\[
 aL_i(t) = aL_i(t-1) + \Delta L_i(t)
\]  

(1)

where \( aL_i(t-1) \) is the activation of letter node \( L_i \) at cycle \( t - 1 \), and \( \Delta L_i(t) \) corresponds to the total activation received by letter node \( L_i \) at cycle \( t \). For \( t = 1 \), \( aL_i(t - 1) \) is equal to zero.

The computation of \( \Delta L_i(t) \) depends on the net activation, \( netL_i(t) \), received by letter node \( L_i \) at cycle \( t \), which can be positive or negative, depending on the ratio between excitations and inhibitions.

If \( netL_i(t) > 0 \),

\[
 \Delta L_i(t) = [max_{act} - aL_i(t-1)] \times netL_i(t)
\]  

(2a)

otherwise,

\[
 \Delta L_i(t) = [aL_i(t-1) - min_{act}] \times netL_i(t)
\]  

(2b)

where \( max_{act} \) and \( min_{act} \) are, respectively, the maximum and minimum activation values of letter nodes. These values are the same as those in the original model—that is, \( max_{act} = 1 \), and \( min_{act} = -2 \). It is this \( \Delta L_i(t) \) transformation that guarantees the nonlinearity in the activation curves.

The computation of \( netL_i(t) \) is determined by the excitations and the inhibitions that are received by each letter node and is equal to:

\[
 netL_i(t) = E_{FL}(t) - I_{FL}(t) - I_{LL}(t)
\]  

(3)

where \( E_{FL}(t) \) and \( I_{FL}(t) \) are, respectively, the excitation and the inhibition received by letter \( L_i \) from the feature level, and \( I_{LL}(t) \) is the lateral inhibition received from the other activated letter nodes. They are calculated in the following way:

\[
 E_{FL}(t) = aFL \times \sum_j aF_j(t)
\]  

(3a)

\[
 I_{FL}(t) = \gamma FL \times \sum_j aF_j(t)
\]  

(3b)

\[
 I_{LL}(t) = \gamma LL \times \sum_v aL_v(t-1)
\]  

(3c)

where \( aFL \) and \( \gamma FL \) are, respectively, the weights of the feature-to-letter excitatory and inhibitory connections, \( \gamma LL \) is the weight of the inhibitory letter-to-letter connections, \( aF_j(t) \) is the activation at cycle \( t \) of feature \( j \) (which is part of letter \( L_j \)), \( aF_j(t) \) is the activation at cycle \( t \) of feature \( k \) (which is not part of letter \( L_j \)), and \( aL_v(t - 1) \) is the activation at cycle \( t - 1 \) of letter \( v \) (being different from \( i \), i.e., any letter other than \( i \) that is activated at cycle \( t - 1 \)). Note that in the IA model both the presence and absence of a feature are used to activate letter nodes. Therefore, when letter \( F \) is presented, for example, the absence of a bottom horizontal line is also used to activate \( F \) and to inhibit letter \( E \). That is, the absent feature number \( 8 \) (see Figure 2) sends a constant activation \((=\alpha FL)\) to letter node \( F \) and a constant inhibition \((=\gamma FL)\) to letter node \( E \).

Contrary to previous versions of the IA model, the feature nodes were able to vary due to the excitatory feedback coming from the letter level. The equations used for feature activation were the same as those for letters. At each processing cycle \( t > 1 \), activation of feature node \( F_i \) is equal to:

\[
 aF_i(t) = aF_i(t-1) + \Delta F_i(t)
\]  

(4)

where \( aF_i(t - 1) \) is the activation of feature \( i \) at cycle \( t - 1 \), and \( \Delta F_i(t) \) corresponds to the total activation received by feature node \( F_i \) at cycle \( t \). This total activation is equal to:

\[
 \Delta F_i(t) = [max_{act} aF_i(t-1)] \times E_{LF}(t)
\]  

(5)

where \( max_{act} \) is the maximum activation value that a feature node can reach. This parameter is set at 2 in the present simulations. \( E_{LF}(t) \) is the feedback excitation coming from the letter level, and it is equal to:

\[
 E_{LF}(t) = aLF \times \sum_j aL_j(t-1)
\]  

(6)

where \( aLF \) is the weight of the letter-to-feature excitatory connections, and \( aL_j(t - 1) \) is the activation of letter \( j \) at cycle \( t - 1 \), letter \( L_j \) being composed of feature \( F_i \). Note that only nodes for present features were allowed to vary in the simulations.