A Biologically Realistic Cortical Model of Eye Movement Control in Reading

Jakob Heinzle
Institute of Neuroinformatics, University Zürich and Swiss
Federal Institute of Technology Zürich; and Bernstein
Center for Computational Neuroscience, Charité–
Universitätsmedizin, Berlin

Klaus Hepp and Kevan A. C. Martin
Institute of Neuroinformatics, University Zürich and Swiss
Federal Institute of Technology Zürich

Reading is a highly complex task involving a precise integration of vision, attention, saccadic eye movements, and high-level language processing. Although there is a long history of psychological research in reading, it is only recently that imaging studies have identified some neural correlates of reading. Thus, the underlying neural mechanisms of reading are not yet understood. One very practical requirement of reading is that eye movements be precisely controlled and coordinated with the cognitive processes of reading. Here we present a biologically realistic model of the frontal eye fields that simulates the control of eye movements in human readers. The model couples processes of oculomotor control and cognition in a realistic cortical circuit of spiking neurons. A global rule that signals either "reading" or "not reading" switches the network's behavior from reading to scanning. In the case of reading, interaction with a cortical module that processed "words" allowed the network to read efficiently an array of symbols, including skipping of short words. Word processing and saccade buildup were both modeled by a race to threshold. In both reading and scanning, the network produces realistic distributions of fixation times when compared with human data.

Keywords: reading, oculomotor control, eye movements and attention, cortical circuit, neural network

Reading is one of the most important tasks learned by humans. Unlike spoken language, which has much earlier origins, writing seems to have been invented about 5,000 years ago to assist commercial activities. But now literacy is seen as a fundamental skill for all members of our society. Unfortunately, reading skills are acquired far more slowly than speech (Dehaene & Cohen, 2007), and in its expert form, reading undoubtedly requires a high degree of cognitive processing, from the low-level visual processing of letters to the high-level semantic processing of sequences of words. Normally, readers need to move their eyes along the lines of text, which they do in a highly efficient way by making a saccade roughly every 250 ms (for a review, see Rayner, 1998).

The standard neurally inspired model of visual word processing (see, e.g., Dehaene, 2009) invokes cortical processing in the "visual word-form" (VWF) region that forms part of the ventral visual pathway. The representation of text is processed in the temporal, parietal, and frontal areas of cortex concerned with phonological and orthographic spoken and written language (see, e.g., Jackendoff, 2002). Dehaene, Cohen, Sigman, and Vinckier (2005) have proposed an algorithm demonstrating how the visual transformations from visual areas V1, V2, and V4 and the left-hemisphere VWF region should proceed. They have supported their model by studies in patients (Gaillard et al., 2006) and by imaging studies in normal subjects (Vinckier et al., 2007).

However, it is also clear that the precise control of eye movements is an essential attribute of normal reading, and thus, in addition to the ventral stream, the dorsal processing stream is also involved. Over decades psychologists have studied in detail the eye movements during reading in an effort to infer the underlying cognitive processes (for a review, see Rayner, 1998, 2009). Several high-level, functional models for the guidance of eye movements have emerged from this line of research (Engbert, Nuthmann, Richter, & Kliegl, 2005; Feng, 2006; Legge, Klitz, & Tjan, 1997; McDonald, Carpenter, & Shillcock, 2005; Pollatsek, Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2006; Reichle, Rayner, & Pollatsek, 2003; Reilly & Radach, 2006; Richter, Engbert, & Kliegl, 2006; Yang, 2006). A crucial component in all these models is the interaction of eye movement control with word processing. These models can accurately reproduce saccade timing and fixation distributions of reading. Nonetheless, it is still not known how the real cortical network, which relies on the interaction of spiking neurons, controls eye movements. Computational models that are biologically realistic can shed some light on how the brain controls reading eye movements and whether the psychological models proposed are feasible. The main goal in our attempt to adapt a realistic model of the frontal eye field (FEF)

area of prefrontal cortex (Heinzle, Hepp, & Martin, 2007) to reading was not to optimize the circuit with respect to a particular task, but rather to show that several tasks can be implemented quite naturally by using the layered structure of neocortex as a computational framework.

The FEF is the key area for the cortical control of saccadic eye movements and attentional signals. In the FEF of monkeys, the first visual responses arise around 60 ms after stimulus onset (Pouget, Emeric, Stuphorn, Reis, & Schall, 2005; Schmolesky et al., 1998), and because FEF is directly connected to extrastriate visual areas (Schall, Morel, King, & Bullier, 1995), this early activity in FEF may modulate visual processing through those areas. In the monkey, microstimulation of FEF with intensities too low to evoke saccades modified the discharge properties of neurons in V4 with corresponding receptive fields and generated attention-like changes in receptive fields and in behavior (Armstrong, Fitzgerald, & Moore, 2006; Moore & Fallah, 2004). In humans, improvement of visual cortical processing was evoked by transcranial magnetic stimulation of the FEF region (Neggers et al., 2007). Several studies suggest that the human FEF causally influences the visual activity in posterior cortical areas by directing overt eye movements and shifts of attention (Deubel & Schneider, 1996; Kowler, Anderson, Dosher, & Blaser, 1995; Ruff et al., 2006; Silvanto, Lavie, & Walsh, 2006; Taylor, Nobre, & Rushworth, 2007).

In humans, FEF is localized in the posterior middle frontal gyrus, rostral to the junction of the precentral and superior frontal gyrus (Neggers et al., 2007), and its likely importance for reading is also indicated by the activation of sensorimotor areas such as the dorsal precentral gyrus during letter and word processing without eye movements (James & Gauthier, 2006; Kujala et al., 2007).

The local neural circuitry, which underlies all cortical computation, consists of a layered network of neurons with strong recurrent synaptic connections (for a review, see Douglas & Martin, 2004). This so-called canonical microcircuit has been used to model the primate FEF (Heinzle et al., 2007), which is crucially involved in guiding voluntary eye movements. The question arises whether the same fundamental circuit could also be the basis of a model of eye movement control in reading. In particular, the model should guide eye movements from left to right and should be coupled to a word recognition process to allow the integration of attention, eye movements, and word processing. It is important to note that this FEF for reading should closely resemble the basic primate FEF circuit for general-purpose visuosaccadic control, as it is very unlikely that a new cortical circuit would have evolved during the short history of reading in humans (Dehaene & Cohen, 2007).

The model presented here is a biologically realistic layered network of spiking neurons that can control sequences of eye movements as in reading. In particular, it proposes a mechanism of how the visual-to-oculomotor transformation can be influenced by a rule, such as the left-to-right bias in reading, and it suggests how eye movements and attention could interact with cognitive processes. Here lexical processing was separated from the oculomotor control and was carried out in a different cortical module, where the length of a word determined the duration that was needed to process it.

The model makes a direct link between high-level, behaviorally realistic models of oculomotor control in reading and the biophys-ical components of spiking neurons and synapses that actually are at the basis of such cortical computations.

Model Description

The local circuit model presented here simulates the layered structure of neocortex. The architecture of the oculomotor part (FEF) follows that presented in a model of the primate FEF (Heinzle et al., 2007). In our previous work, we demonstrated how a "canonical" cortical circuit for oculomotor control could flexibly interact with a cortical module that recognized shapes of objects to guide different voluntary eye movements, such as normal saccades, antisaccades, and delayed-memory saccades. The FEF model used here has the same architecture with only two changes. First, the internal connection that allowed the network to guide eye movements according to a rule was set to reflect the left-to-right bias in Western culture (Spalek & Hammad, 2005). Second, the FEF interacted with a second cortical module, a word processing (WP) module that was also modeled with spiking neurons. The WP module processed words according to their length (corresponding to difficulty) and directly interacted with attentional control in the FEF. The VWF region in the temporal cortex (Cohen, Dehaene, Vinckier, Jobert, & Montavont, 2008; Glezer, Jiang, & Riesenhuber, 2009; Vinckier et al., 2007) could constitute such a WP module. Note that the focus of the FEF model is on the cortical control of eye movements, not on a precise modeling of the processing of written language (Dehaene et al., 2005).

The functional architecture of the model is presented in the following sections. First, the generic, oculomotor part of the network is described. Then the modifications to the FEF model that were sufficient to introduce reading behavior in the circuit are explained. Crucially, these modifications did not interfere in any way with the functions of the original circuit. For simplicity, we focus here on the parts that are important for reading. The adaptations introduced new functions to the original network by exploiting the ability of gating different functional pathways by neuronal inputs.

A full mathematical description of the network model, including the dynamics of the integrate-and-fire (IF) neurons and synapses and of the pattern of synaptic connections, is given in the Appendix. It is important to bear in mind that the dynamics of the network resulted directly from the recurrent interaction of spiking (IF) neurons connected by simple exponential synapses and not from biophysical peculiarities of individual neurons or complicated synaptic dynamics.

A Model for Eye Movements During Reading

Reading requires a tight interaction of word processing and oculomotor control. Current models of eye movement control are divided into two classes. Serial lexical processing models such as the E-Z Reader model (Pollatsek et al., 2006; Rayner, Juhasz, & Brown, 2007; Reichle et al., 1998, 2006, 2003) process one word after the other and jump to the next word only when the last one is recognized. Others have proposed a parallel lexical processing model, such as SWIFT (Engbert, Longtin, & Kliegl, 2002; Engbert et al., 2005; Richter et al., 2006), whereby several words are processed at the same time. The eye movement strategy of our model followed the idea of serial processing, and hence the atten-

tional focus was sequentially scanning one word after the other and moved only once the attended word was sufficiently processed to allow the attentional shift. Although the behavior is serial, the recurrent nature of the network model makes it impossible to infer on the basis of the cortical connection a hierarchical differentiation between lexical processing and oculomotor control. The two processes constantly interact and influence each other.

To be able to model "reading" directly with a neural network, we introduced a simplified language consisting of three types of "words" (short, medium, and long) only. Words were represented by sequences of *x*s (short: *x*; medium: *xx*; long: *xxx*; see Reichle & Laurent, 2006, for a similar approach), and the length of a word was taken as the parameter defining how difficult a word is to recognize. Figure 1 shows the required interaction of attention, eye movements, and the recognition of words that defined the set of permitted reading saccades for our model. After each saccade, the focus of attention should again be at the fovea. When the foveated word was recognized (i.e., sufficiently processed), attention was free to shift to the next target. Once attention was shifted to a nonfoveal word, two processes started in parallel—motor buildup (in the FEF) and word processing (in the WP module)—and the outcome of the race between these two decided the next behavior. If the saccadic buildup was faster, a normal saccade was triggered, and the same sequence began again. If, however, word processing was faster than the saccade preparation, attention was shifted a second time, and a new saccade was planned. This second case corresponded to the skipping of a word. A similar skipping mechanism was proposed for the E-Z Reader model (Reichle et al., 1998). As long as there were words to the right of the position currently attended to, the network shifted its attention to the right.
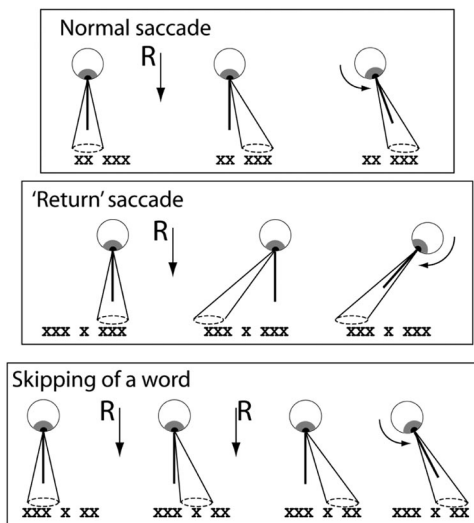

Figure 1. Rules for "correct reading" of the model: schematic drawing of the interaction of attention (indicated by the angle and dashed circle), eye movements (indicated by the curved arrow), and word recognition (indicated by *R* and the vertical arrows) during the three basic types of reading eye movements. The direction of gaze (fovea) of the eye is shown by the bold line. Note that every saccade is preceded by at least one attentional shift. During the skipping of a word, the short word *x* is processed in the parafoveal region, and attention is shifted a second time before the saccade is made.

Only at the end of the line, a large saccade to the left needed to be made back to the beginning of the line.

The Basic FEF Circuit for Oculomotor Control

The control of eye movements requires several computational steps: the visual selection of saccade targets, the allocation of attention, and, of course, the motor output that drives the eye movements. A canonical cortical circuit of the FEF was proposed that incorporated these steps of processing. We briefly summarize the functioning of that circuit here. A detailed description of the FEF microcircuit is given in the Appendix (see also Heinzle et al., 2007).

The layers of the FEF were related to the specific stages of the visual-to-oculomotor transformation. Layer 4 received a dorsal visual input from early visual areas. It selected the retinotopic position of the strongest of those. Layer 2/3 transformed the phasic visual signal of Layer 4 into an attentional signal. It connected to a WP module (which represented the ventral cortical stream) and activated a word recognition mechanism at the currently attended retinotopic position. The WP module processed a feature-specific input from the ventral stream and sent a signal back to Layer 2/3 of the FEF when a target was recognized. Recognition in this context meant that the attentional focus could be moved from the current position. In addition, attentional activation in Layer 2/3 drove the motor neurons[1] in Layer 5. Therefore, Layer 2/3 signaled the focus of covert attention as well as the motor plan for the next saccade. Layer 5 consisted of two parts: "ramping" motor neurons that built up activity (L5r) and "burst" motor neurons (L5b) that signaled the motor output to the superior colliculus and the brainstem.

The selection of the next target during reading followed a rule. Neurons in Layer 6 projected back to Layer 4 and biased the visual selection. They provided a top-down saliency that influenced the visual selection in Layer 4 according to the currently attended location. In the present model, this top-down bias provided the left-to-right bias for reading (see also Figure 3A). Neurons in Layer 6 were either driven by Layer 2/3 and had attention-related activity (L6r) or driven from Layer 5b and showed saccadic bursting activity (L6s).

Figure 2A sketches the layers of the network and explains their role in the visual-to-oculomotor transformation. Only feed-forward connections within the network are shown. These connections represent the general flow of information but are not necessarily the strongest connections in the circuit.

The functions described above were mapped onto a one-dimensional topographic layered network that spanned the horizontal axis of the visual field with populations of neurons at 21 equally spaced retinotopic positions. The anatomy of the network (cf. Figure A1) respected the cortical anatomy, and connections were inserted in the network according to the connectivity pattern known from cat visual cortex (Binzegger, Douglas, & Martin, 2004). Figure 2B shows the layered, retinotopic network structure.

---

[1] In the literature, neurons that show a strong response around the time of the eye movement are sometimes called premotor, saccadic, presaccadic, movement, and motor neurons. We use the term *motor neurons* consistently throughout this article, which is not to be confused with motor neurons in the brainstem and superior colliculus or motoneurons of the eye.
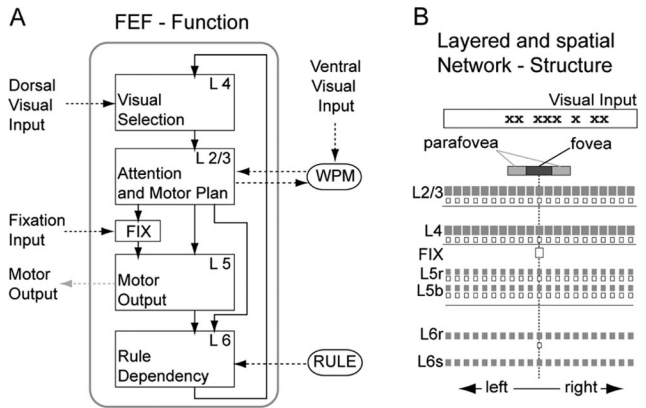
Figure 2. Functional description of the model. (A) Each layer within the frontal eye field (FEF) was assigned a distinct function in the visual-to-oculomotor transformation. Connections that are shown indicate the general flow of information within the network. Layer 2/3 of the FEF interacted with the word processing module (WPM). Layer 6 of the FEF was targeted by a rule input that signaled to the FEF whether it should read. The fixation (FIX) input directly targeted the fixation population and therefore enabled an external fixation signal to inhibit motor buildup. Although fixation is a necessary component of the wider functions of the FEF circuit, the fixation-specific input was not needed as a control parameter in the reading context. (B) The FEF model represented space along the horizontal axis, as indicated. The foveal and parafoveal region of the network for the different layers are shown with respect to the retinotopic representation of the FEF.

Reading Bias and Interaction With Word Processing

The FEF circuit model included two important pathways that allowed it to have a high flexibility for its oculomotor control. The "rule input" from Layer 6 to Layer 4 provided a top-down bias that could influence the visual selection in an attention- and saccade-dependent way. This connection was able to bias visual selections to the right of the fovea as in reading. The FEF interacted with the WP module to guarantee that attention was withdrawn from a retinotopic position only when the stimulus at that position was processed. We describe below how these two mechanisms were adapted to guide eye movements that followed the reading strategy described above.

Reading Bias in the Visual Selection

The left-to-right bias in the eye movement pattern, which is obvious during reading, can also be observed in a reaction time task that investigates effects of inhibition of return. Hence, although the direction of the bias depends on the direction of reading, the bias is not directly related to reading (Spalek & Hammad, 2005). In the FEF model, the left-to-right bias was introduced by the connection from attention-related neurons in Layer 6r to excitatory neurons in Layer 4 (see Figure 3A). When active, this input slightly enhanced the excitatory drive to neurons in Layer 4 that represented positions to the right of the position currently attended to. An input from the dorsal visual stream at those positions was more likely to be selected than one at an unbiased position. An even weaker excitation targeted the positions on the far left of the visual field. This weaker bias made the

network generate a saccade back to the beginning of a line when there was no word to the right of the fovea. It is important to note that we assumed the visual reading rule to contain both the rightward and the far-left bias simultaneously. Nevertheless, as long as there were words to the right of the fovea, the stronger bias was
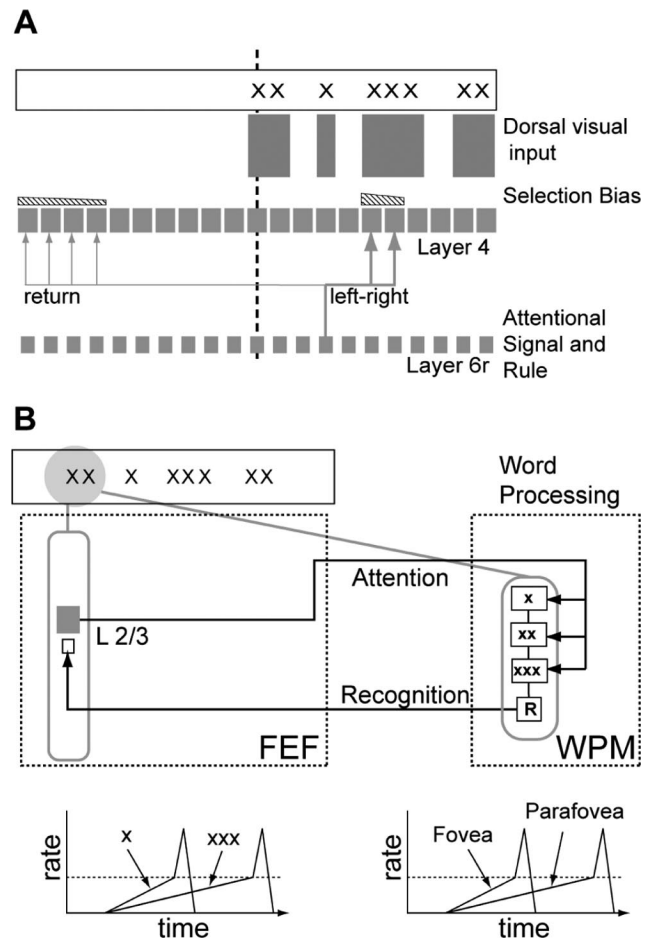

Figure 3. Implementation of the reading-specific rules. (A) Left-to-right bias in reading. Visual selection in Layer 4 was biased by the input from Layer 6r. This input was not strong enough to drive Layer 4 (textured boxes above layer indicate its strength). However, it facilitated the selection of the visual input to the right (left–right bias) and to the far left (return bias). (B) Interaction of frontal eye field (FEF) and word processing module (WPM). The WPM was activated only at the attended position, given by the input from Layer 2/3 of the FEF (see also Connection R1 in Figures A1 and A2). At this position, the visual input representing the ventral cortical stream drove word processing according to the length of the attended word. Only one retinotopic position is shown. The speed of its recognition depended on the length of the word (see bottom, left). Short words were processed faster than long words. In addition, the time needed to process a word depended on the retinotopic position of the attended target. At the fovea, word recognition was fastest; in the parafovea, it was considerably slower; and outside the parafovea, words could not be processed (see bottom, right). The black lines indicate neural firing in the WPM. Note the ramping and the burst after threshold crossing (dashed horizontal line). The output of the WP module network targeted the inhibitory neurons in Layer 2/3 of the FEF and suppressed attentional activity after a word was sufficiently processed (Connection R2 in Figures A1 and A2).

effective. In line with the small rule-dependent changes observed in monkey prefrontal cortex (Everling & DeSouza, 2005), the activity in the rule–bias connection was weak. It could not directly drive visual selection, but it did result in increased spontaneous activity in Layer 4.

Ramping of FEF Layer 5r and Word Processing

The neurons in Layer 2/3 drove ramping neurons in Layer 5r. These neuronal populations slowly increased their firing depending on their input. The time the ramping neurons needed to build up enough activity to trigger a saccadic burst varied due to the unspecific random inputs to the network that mimicked background cortical activity. However, this timing did not depend directly on any visual feature. Lexical processing, by contrast, should depend on the difficulty (or length in the case of our model) of the word to be recognized. This was simulated by a similar ramping-to-threshold behavior in the WP module. Ramping neurons in the WP module started integrating the visual input at the position signaled by Layer 2/3. Hence, attentional interactions of the FEF with the WP module originated from Layer 2/3, whereas the motor output resulted from Layer 5. The same layer specificity has been observed in anatomical studies in the monkey (Barone, Batardiere, Knoblauch, & Kennedy, 2000; Pouget et al., 2009). When the ramping neurons reached threshold, a population of bursting neurons was activated, and these excited inhibitory neurons in Layer 2/3 of the FEF. This inhibition released attention from the position currently attended to. Hence, word recognition in this context meant that the word in the attentional focus was processed sufficiently for attention to shift away from it. The slope of the ramping defined the speed of word processing and was controlled by the input to recognition-ramping neurons, which in turn depended on the word length (see Figure 3B, bottom left).

In keeping with earlier research on reading (Rayner, 1998), words were processed within two distinct spatial regions at different speeds (compare Figures 2B and 3B). First, the foveal region was defined to contain the three central positions of the retinotopic array. Words that fell within this region were recognized rather quickly. Second, the parafoveal region contained two positions either side of the fovea. Words in the parafoveal region were recognized more slowly (see Figure 3B, bottom right). Because the retinotopy of the FEF model was not a linear metric map of the visual field, the ranges of the fovea and parafovea cannot directly be compared with the physiological range of 1° to both sides of the fixation (around six to eight letters of normal script) for the fovea and 5° for the parafovea.

The race between motor and word recognition buildup in the parafoveal regions controlled word skipping. If the saccadic burst was produced first, a normal saccade to the next word was made. The visual input was updated accordingly, attentional activity was internally reset to the foveal position in Layer 2/3, and word processing continued. If, however, the word was recognized first, the attentional activity in Layer 2/3 was suppressed, which stopped the buildup in Layer 5r. In this case, attention shifted further to the right, and a word was skipped.

Simulations of Reading and Scanning Tasks

The network's behavior was tested in the reading task. Ten versions of the network, differing in the random connections and

weights, were simulated over a period of 100 s while scanning random "text." The text input to network consisted of an array of words that spanned maximally 11 retinotopic positions (e.g., from Positions 0 to 10 when the fovea was at the beginning of the line). This size limitation assured that the network could always see the complete line. Whenever the network made a saccade back to the first word of the line, the word pattern of the line was changed randomly. In a second set of simulations, the performance of the network was assessed while the network read the same line again and again. The network scanned the same line 10 times for a period of 10 s, which resulted in 100 s of simulated time for each line.

Eye movement statistics were assessed for all simulations directly from the firing of Layer 5b. In particular, we examined three measures and how they depended on word length. *Average fixation duration* was defined as the mean of all fixations on a particular class of words. There was no distinction between saccades to the right or to the left and according to whether the word was fixated a second time. *Total fixation duration* was defined as the sum of all fixations that occurred on a single word.[2] Words that were never fixated did not enter in this analysis. These fixation measures were averaged over all words corresponding to the same class. The third measure, *probability of skipping*, was defined as the probability that a word was never fixated.

We then checked how many of the saccades followed the reading rules. Saccades were evaluated according to the criteria illustrated in Figure 1. Normal forward saccades had to land on the next word or could start on the first letter of a long word and end on the third letter (double fixations). Because the reading bias included the Positions 2 and 3 to the right of the fovea, such double fixations could occur on long words only. Skippings of words were correct only if the word that was skipped was correctly recognized by the WP module. (In experiments with human readers, it is impossible to assess for certain whether a skipped word was processed by just examining eye movements.) Return saccades were correct if they ended on the first word of the line and had started on the last word. Some return saccades that did not start on the last word of the line, but correctly skipped and processed the last word, were also considered to be correct. All other saccades were counted as errors. For example, because of spontaneous fluctuations in Layer 2/3, some attentional shifts occurred without the recognition of the word. It is important to note that the WP module did not include any semantic processing that would allow the network to make regressions within one line. Hence, all saccades to the left that did not end on the first word of a line and therefore were not return saccades can be regarded only as errors.

The network was also run with the rule input switched to "scanning" mode. In this case, the network freely scanned an array of six targets that differed in their intensity without any left-to-

---

[2] A number of different eye movement measures are typically used including *first fixation duration*, which is the duration of the first fixation on a word independent of whether it is a single fixation or the first of multiple fixations; *single fixation duration*, which is only one fixation on a word; and *gaze duration*, the sum of all fixations on a word before moving to another word. All these measures are calculated for fixations during the first read through the text. In the present analysis, we considered only average fixation duration and total fixation duration. We calculated these two measures for all fixations of a word. Words that were never fixated did not enter the calculation of fixation times.

right bias in the selection of visual targets. To demonstrate the effect of different strengths (e.g., luminance or contrast) of visual inputs, we assigned different relative strengths to the six targets (1, 0.9, 0.8, 0.9, 1, 0.8; see Figure 10). The task of the network was to look freely around in the visual scene over the time of 100 s. Note that the specific strength of the inputs was not critical for the scanning, but the chosen values nicely illustrate the network's behavior. Finally, we tested the network in the absence of any visual input. Simulations were run in MATLAB (MathWorks, Natick, MA). The code[3] for the FEF model of reading was inspired by Salinas (2003).

Model Simulations

Simulations of the FEF model circuit showed that indeed the network was able to produce reading-like eye movements when presented with lines of simplified written text (which consisted only of three words of different length). Here we first present the dynamics of the network during single saccades and show the relation of the firing pattern of neurons in specific layers to saccades, attention, and word processing. Then we give a detailed description of the eye movement statistics and the interaction of attention, eye movements, and word processing. Finally, we tested the behavior of the network for scanning without the reading rule.

Single Saccades During Reading

The interaction of spiking neurons in the layered network implemented the dynamics of the network. The firing of single neurons was closely related to the eye movements. Figure 4 shows the spike raster plots and the spike histograms for three excitatory neurons in Layers 4, 2/3, and 5. The spiking of a typical single neuron for many identical saccades ($n = 146$) into its receptive field is compared with the average firing of the corresponding population of neurons. Temporal patterns of visual neurons are shown with respect to the visual onset (i.e., the last saccade), and motor neuron activity is aligned to the actual saccade. Neurons in Layer 4 had visual activity that started around 50 ms (Schmolesky et al., 1998) after the last saccade, and their activity was suppressed before the upcoming saccade. Neurons in Layer 2/3 showed sustained firing until the time of the next saccade. These neurons signaled the attention and the motor plan for the next saccade. Neurons in Layer 5r slowly increased their firing toward the time of the saccade. The slope of the ramp in these neurons defined the time needed for motor planning. Groups of neurons with similar connectivity patterns formed the functional units of the FEF network. These groups showed average population firing rates similar to those of the single neurons (see Figure 4, right panels). The similarity of the two was a result of the similar connection patterns and, therefore, the similar inputs received by neurons within the same population. A detailed discussion of the spiking behavior of all neuron types within the model FEF is given in Heinzle et al. (2007).

The functioning of the FEF circuit and its relation to behavior is best described by looking at the population activities of excitatory neurons in all layers and the firing of the WP module. Inhibitory neurons were not as closely linked to attention, eye movements, and word processing, but instead modulated the activity of excitatory neurons and were involved in resetting the network after
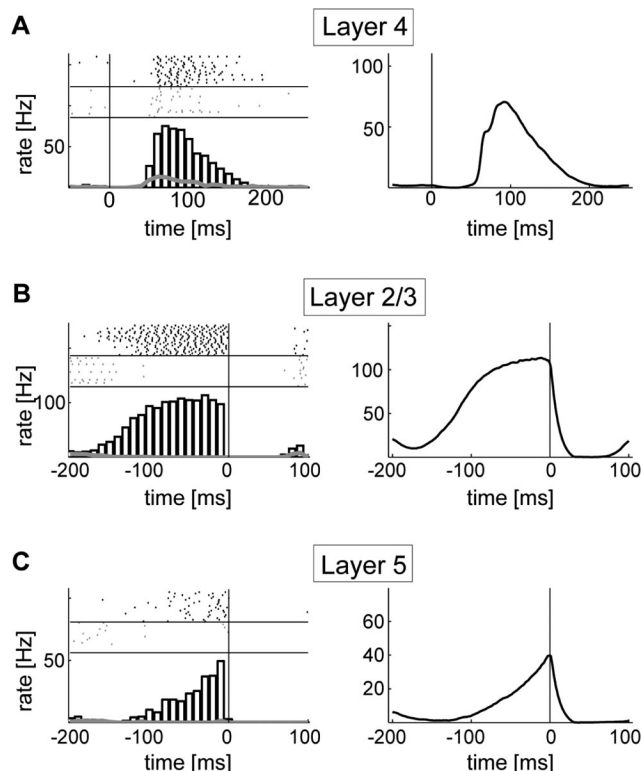

Figure 4. Single-cell firing and population rates. The firing of typical neurons (left panel) is illustrated by the spike raster plot for 20 saccades into the receptive field of the neuron (black dots) and 20 saccades to a different position (gray dots). The spike histograms below show the average firing rate of the neuron for 146 saccades into (bars) and 209 saccades outside its receptive field (gray line). The average population activity of all neurons with the same receptive field (right panel; ordinate is average spike rate in hertz) was similar to the response of a single neuron of that population. (A) Neuron and population in Layer 4. Trials are aligned to the onset of the visual input, equivalent with the time of the saccade (vertical line) that preceded the visual input by 50 ms. (B) Neuron and population activity in Layer 2/3. Trials are aligned to the motor response, that is, the time of the saccade into the receptive field of the neuron (vertical line). (C) Neuron and population in Layer 5r. Same timing as in Figure 4B.

every saccade. The firing of populations of excitatory neurons during the three basic types of saccades (see Figure 1) is presented here. Figure 5 shows the neural activity for a normal forward (left-to-right) saccade during reading. The firing of populations of neurons was directly related to processing steps during the saccade. Immediately after the last saccade, the foveal population of Layer 2/3 (gray trace) was activated due to internal resetting of activity and signaled attention to be at the fovea (Time Point A) until the word was recognized. The recognition of a word was signaled by the burst neurons in the WP module (Time Point B). Here recognition did not correspond to complete lexical access of a word, but rather indicated that the word was processed sufficiently for visual attention to be withdrawn from it. Fifty millisec-
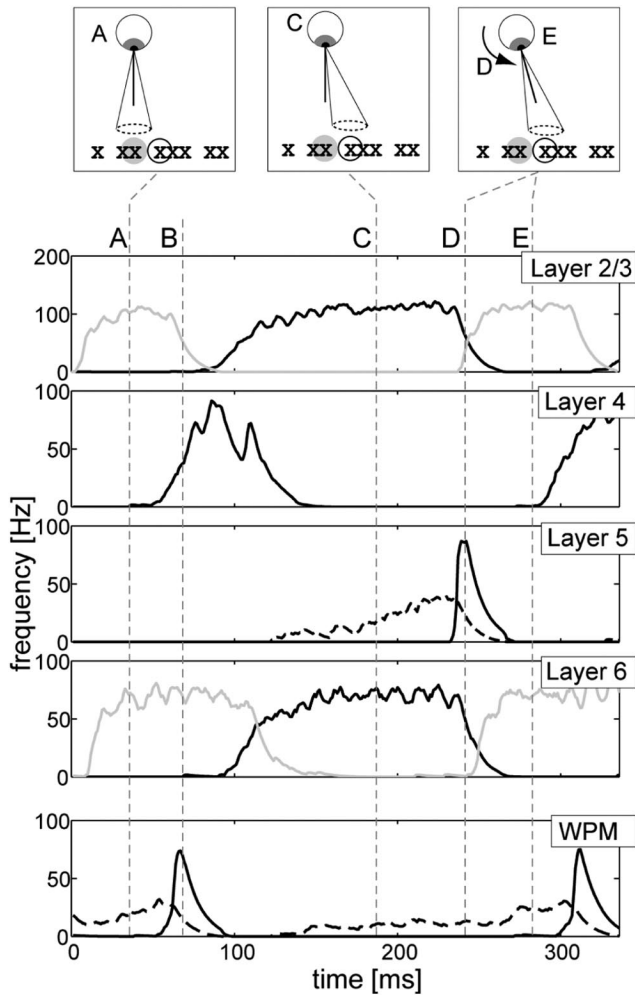
Figure 5. Activation of populations of excitatory neurons for a normal saccade. Gray traces correspond to neurons representing the fovea, and black traces correspond to neurons representing the saccade target (see schematics above). All traces are aligned to the previous saccade. Note that visual input reached the frontal eye field 50 ms later. Schematics show the corresponding eye movement (cf. Figure 1). (A) After the saccade, attention is on the fovea (Layer 2/3), and the two-letter word is processed in the word processing module (WPM). (B) The WPM signals the recognition of the word. Attention is released from the fovea, and Layer 4 selects the next target. (C) Attention is on the first *x* of the three-letter word. Activity is ramping up in Layer 5r (dashed trace) as well as in the WPM (dashed trace). (D) A saccade is triggered by the burst in Layer 5b. Attention is immediately shifted back to the fovea, which is now shifted to the new position. (E) Attention remains on the fovea while the word is still being processed.

onds after the saccade, the new visual input (in the postsaccadic retinotopic coordinates) reached Layer 4 of the FEF. The input to Layer 4 consisted of a dorsal visual input plus the bias input from Layer 6r. Layer 4 selected the strongest input, even if attention was still needed at the fovea. Once the word at the fovea was processed, Layer 2/3 was suppressed at the attended location (gray trace), and its activation shifted to the position selected by Layer 4 (black trace). Layer 2/3 directly drove ramping neurons in Layer

5r and, at the same time, activated the recognition process in the WP module (Time Point C). It is important to note that the processing of the word that is the target of the next saccade has already started at this point. The buildup activity of ramping neurons in Layer 5 (dashed trace) reached threshold and triggered a saccade (Time Point D). After the saccade, Layer 2/3 was activated at the fovea again, and the network was reset to almost the same state as at Time 0 (Time Point E).

Large return (right-to-left) saccades back to the beginning of the line followed a similar pattern of activation. Figure 6 shows the activity of the FEF and WP module during such a return saccade. Unlike with a small forward saccade, the WP module was not able to start processing the word at the target position, as it was outside the parafoveal region. Hence, word recognition could start only after the saccade was made.

In contrast to the single saccade, the skipping of a word required two shifts of attention (see Figure 1). A typical activation pattern
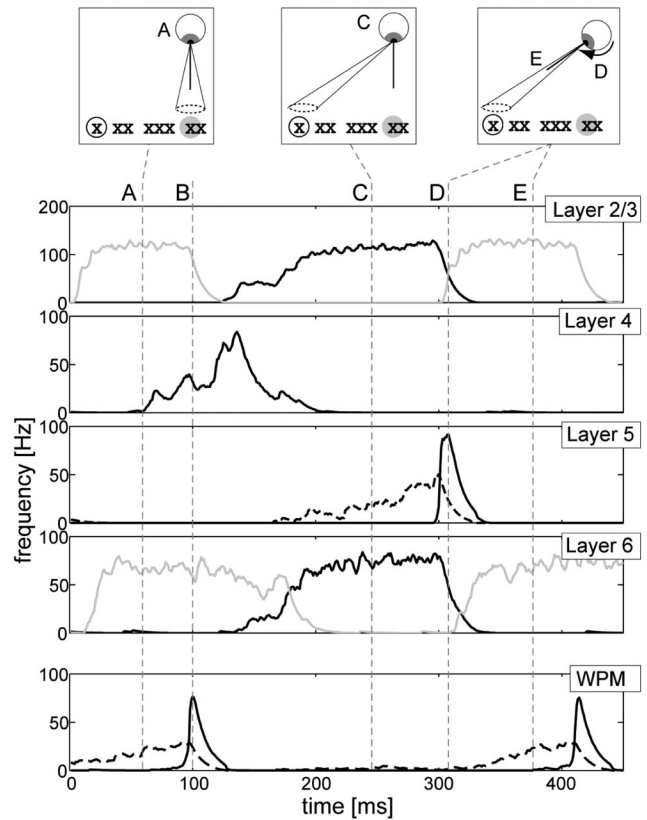
Figure 6. Activation of populations of excitatory neurons for a return saccade. (Same conventions as in Figure 5.) (A) After the saccade, attention is on the fovea (Layer 2/3), and the two-letter word is processed in the word processing module (WPM). (B) The WPM signals the recognition of the word, attention is released from the fovea, and Layer 4 selects the next target. (C) Attention is allocated to the first word of the line. Activity ramps up in Layer 5r (dashed trace). Note that the WPM cannot process the word because it is outside the parafoveal region. Hence, there is no buildup activity related to word processing (dashed line). (D) The saccade is triggered by the burst in Layer 5b, and attention is immediately shifted back to the fovea, which is now shifted to the new position. (E) Attention remains on the fovea, and activity in the WPM now builds up to recognize the word (dashed line).

of the network during skipping is shown in Figure 7. As for the normal saccade, the target was selected and attended to (Time Point C). However, the buildup of activity in the WP module reached threshold before the ramping in Layer 5r triggered a saccadic burst (Time Point D). The recognition burst suppressed activity in Layer 2/3 and allowed the network to shift the focus of attention further to the right (Time Point E). After this second shift of attention, the target word lay outside the parafoveal region. Therefore, the WP module did not build up activity, and the saccade was triggered by the burst in Layer 5 neurons (Time Point F). As soon as the attentional Layer 2/3 was activated, Layer 4 neurons at the same retinotopic position were suppressed, so that the second target could be selected while the first target was still

Figure 7. Activation of populations of neurons in the frontal eye field during the skipping of a word. (Same conventions as in Figure 5; black dotted traces correspond to neurons representing the skipped position.) (A) After the saccade, attention is on the fovea (Layer 2/3), and the three-letter word is processed in the word processing module (WPM). (B) The WPM signals the recognition of the word, attention is released from the fovea, and Layer 4 selects the next target (dotted line). (C) Attention is on the first short word, and activity ramps up in Layer 5r (black dotted line) as well as in the WPM (dashed line). (D) The target is recognized by the WPM (solid line). In Layer 2/3 (black solid line), attention is shifted to the second short word. Note that Layer 4 had already selected this target before the first word was recognized. (E) Attention is on the second short word, and activity ramps up in Layer 5r (dark gray trace) but not in the WPM (dashed line). The target word is outside the parafovea. (F) A saccade is triggered by the burst in Layer 5b (black solid line). Attention is immediately shifted back to the fovea. (G) Attention remains on the fovea while the second short word is processed by the WPM (dashed line).

being attended (Time Point C). This prospective visual selection allowed the network to change its focus of attention quickly once the attended word was sufficiently processed.

Reading Statistics

The reading behavior of the network was assessed by examining the interaction of eye movements, attention, and word processing. This first part focuses on the statistics of eye movements and how they compare with those observed in humans. It is clear that the model, because of its limited spatial resolution and the highly simplified lexical processing mechanism, was not able to fully reproduce the detailed pattern of eye movements. Nevertheless, its behavior was comparable in some key qualitative aspects to experimental findings.

The network was run 10 times for 100 s on different random texts. Figure 8A shows an example fixation pattern on five lines of text. Recall that the text length of the lines was such that the network could always see the whole line. A detailed list of all saccades shown in Figure 8A is given in Table 1. During the 1,000 s of reading, 3,569 saccades were made with an average fixation time of 278 ± 87 ms. Figure 8B shows the distribution of fixation times, which is comparable to that of human readers (see Rayner, 1998, Figure 1). With respect to the fixation times on words of different length, the following pattern emerged: Although average fixation durations were very similar for the different lengths of words, the total fixation duration was strongly correlated with word length (Figure 8C; compare Reichle et al., 2003, Figure 8, and Pollatsek et al., 2006, Figure 6, for similar graphs related to word length and word frequency), which was mainly due to double fixations on long words. The spatial pattern of saccade vectors is illustrated in Figure 8D. The reading bias (see Figure 3A) is clearly reflected in this distribution by the 87% of the saccades made into the region of the bias. When skipping was also considered, another 8% of the saccades were made into the region that was within the bias. Two aspects in the pattern of saccades to the left (with negative vector) should be noted. Saccades to Positions −2 to −6 did not fall within the bias and were mostly errors, whereas the saccades to Positions −7 to −10 where nearly all correct return saccades. Although the return bias was strongest for Position −10, this was not reflected in the saccade distribution, because many lines were shorter than 10 "letters," and therefore the return saccades often could not be of maximal length (Figure 8D). The skipping of words was best characterized by counting the words that were never foveated. Around 30% of the short words were skipped, whereas less than 3% and 1% of medium and long words, respectively, were skipped (Figure 8E). The above definition for word skipping depended only on the fixation pattern, irrespective of whether skipped words were actually processed.

Although the above statistics summarized the global fixation pattern, individual fixation times differed systematically according to several factors. For example, fixation times depended on whether the fixation was followed by the skipping of a word or whether the fixation was the first in the line. The overall statistics did not make any distinction between these different contexts of fixation. To compare in more detail the fixation distribution obtained by our model to existing data, we grouped fixations according to different aspects and then compared fixation times (see Table 2). It is important to note that the model was not fine-tuned
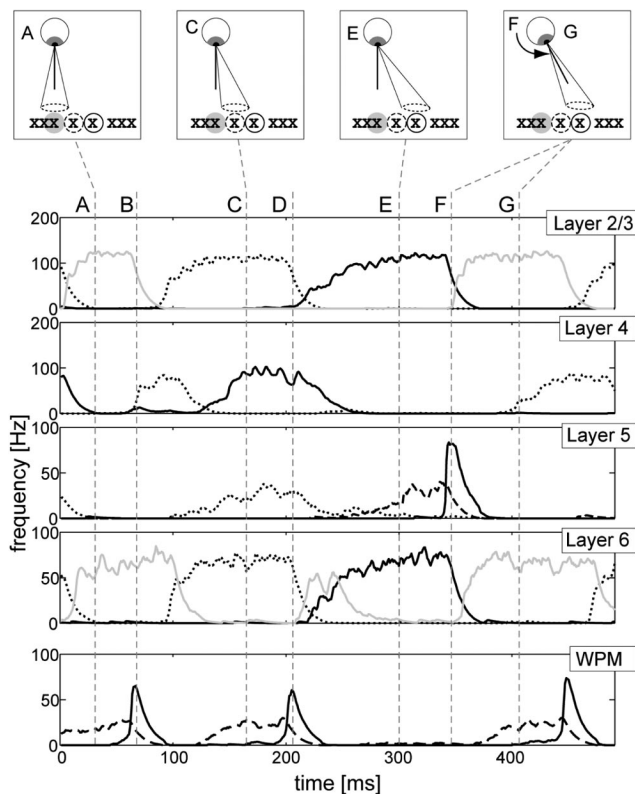
to fit any of the effects discussed below. Rather the specific timing emerged from the design of the model. Table 2 summarizes this more detailed analysis. For example, prior to skipping, fixation times were increased by about 180 ms. A similar but smaller effect was observed in humans (Pollatsek, Rayner, & Balota, 1986; Rayner, Ashby, Pollatsek, & Reichle, 2004). However, the analysis of a large corpus revealed a more subtle interaction that depends on the frequency of the word that is skipped, with in-
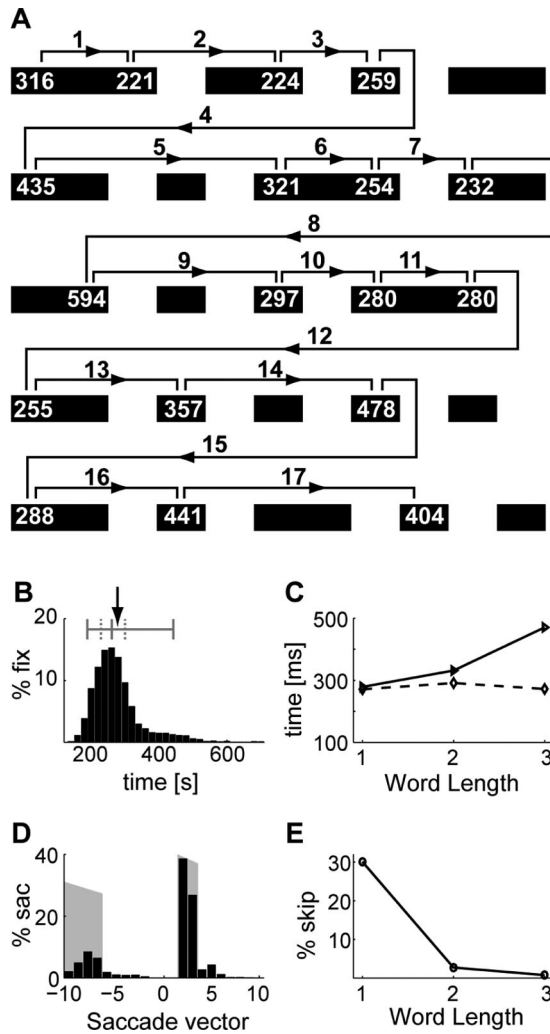
creased fixations prior to long and low-frequency words but no increase prior to short and high-frequency words (Kliegl & Engbert, 2005). Importantly, the total time used to recognize two words—the fixated and the skipped—was shorter than the sum of two single fixations. Hence, skipping was beneficial in terms of reading speed. Although the reprogramming of the saccade took some time, it was faster to switch attention and reprogram the saccade than first to make the saccade and then to start reprogramming. Our model did not increase fixation times in order to finish parafoveal processing, as suggested by a reinforcement learning model (Reichle & Laurent, 2006). Similarly, our model did not include any mechanism that affected the parafoveal preview according to difficulty of the foveated word (Henderson & Ferreira, 1990).

Table 1

List of All Saccades in Figure 8A

Number	Type[a]	Position[b] From	Position[b] To	Fixation time (ms)
1	Double fixation	0	2	316
2	Forward	2	5	221
3	Forward	5	7	224
4	*Return*	7	0	259
5	Skipping forward	0	5	435
6	Double fixation	5	7	321
7	Forward	7	9	254
8	Return	9	1	232
9	Skipping forward	1	5	594
10	Forward	5	7	297
11	Double fixation	7	9	280
12	Return	9	0	280
13	Forward	0	3	255
14	Skipping forward	3	7	357
15	Skipping return	7	0	478
16	Forward	0	3	288
17	*Skipping*	3	8	441
18	Skipping return	8	0	404

[a] Errors are in italics. [b] Starting and landing position of the saccade, where 0 is the first position within a line.

Table 2

Detailed Fixation Time Statistics (ms)

Measure	Short words	Medium words	Long words
All fixations[a]			
All	271 ± 79	291 ± 97	272 ± 82
After skipping	319 ± 111	306 ± 83	374 ± 236
First in line	293 ± 70	319 ± 79	311 ± 62
Forward only[b]			
All	250 ± 46	261 ± 52	249 ± 52
After skipping	276 ± 42	284 ± 30	316 ± 53
First in line	267 ± 31	287 ± 34	299 ± 38
No preview[c]	275 ± 33	289 ± 43	307 ± 35
Double fix[d]			261 ± 44
Prior to skip[e]	428 ± 49	440 ± 60	434 ± 46

[a] All fixations, irrespective of the saccade following the fixation. [b] Considering only fixations that were followed by a normal forward saccade to the next word. [c] Results from test simulation without parafoveal preview. [d] First fixations in double fixations of the same word. [e] Fixations before skipping of a word.

Figure 8. Reading statistics. (A) Sample trace of 18 fixations on five consecutive lines of text presented in the retinotopic frame in which the fovea fixates the first letter of the line. Black numbered arrows indicate saccades. White numbers give the fixation times. Compare Table 1 for a complete list of the 18 saccades. (B) Fixation time distribution for a total reading time of 1,000 s ($N = 3,569$). Gray lines above indicate 5th, 25th, 50th, 75th, and 95th percentiles. The arrow shows the mean fixation (fix) time (278 ± 87 ms). (C) Mean fixation time (diamonds) and total fixation duration per word (triangles) as a function of word length. (D) Saccade (sac) vector distribution. The distribution of the saccade vectors for the 3,569 saccades is given. Note that only very few saccades fell outside the biased region for reading (shown by the gray shading). (E) Probability of skipping a word as a function of word length. All skippings, even without recognition of the skipped word, are included in this graph.

Fixations after return saccades and after skippings were around 30 ms longer than fixations that followed normal forward saccades (see Table 2). In both cases, the parafoveal preview was not active, which resulted in an increased time needed for processing a word at the fovea. The same effect was reported in reading for first fixations in a line (Heller, 1982). When, as a test, the network was run in a setting in which words could be processed only within the fovea, fixation times were similar to those for the first fixations in a line and fixations after skippings. Indeed, in human readers, fixation times increase when the fixated word is visible only and all other words are masked (Rayner, 2009).

Interactions of Attention, Word Processing, and Eye Movement

Because the FEF model included different populations of neurons that signaled attention, word processing, and eye movements, it was possible to compare the timing of these mechanisms directly. Attentional shifts occurred within around 20 ms. This fast reallocation of attention was possible due to the interaction of the visual saliency map in Layer 4 and the attentional map in Layer 2/3. Word recognition times at the fovea and the time needed for saccadic buildup were calculated in the simulations without parafoveal preview. Words were attended at the fovea for 104 ± 11 ms (short words), 121 ± 15 ms (medium words), and 132 ± 15 ms (long words) before the WP module sent the recognition signal. Hence, the word recognition time corresponded approximately to the first stage of lexical processing in the E-Z Reader model (Reichle et al., 2003). Saccadic buildup was independent from word length and resulted in 144 ± 27 ms.

An additional interesting point that requires the exact knowledge of attention and word processing mechanisms concerns the error statistics. Although the distributions of saccades and the fixation times of the network reproduced many aspects of reading, its full performance could be assessed only by looking at every saccade. In particular, we assessed how many of the saccades followed exactly the pattern of attention, word recognition, and eye movements illustrated in Figure 1. For this, the network simulated the scanning of the same line of text for 100 s with several input lines. The activation pattern within the FEF was then checked for every saccade. Figure 9A shows a 10-s trace of the attentional spotlight (dotted line) and the gaze (solid line). Together with the recognition of words, these were used to classify each saccade. An example of an error is illustrated (with an arrow). In this case, an attentional shift occurred without recognition of the word, which resulted in the skipping of two short words.

Figure 9B summarizes the errors for different lines of text. In all lines, over 80% of saccades followed exactly the desired pattern. Most errors occurred in lines with many short words, and during return saccades. The network made most errors on the first line, which consisted of six short words. However, to examine some specific errors in more detail, we consider the fourth and the seventh lines. Recall that the network received noisy input and that any deviation, as, for example, a fluctuation in the attention signal in Layer 2/3 or a double recognition of the same word, was counted as an error.

In the fourth line (*xx x xx x*) of Figure 9B, 46 of 321 saccades (14%) were errors. Twenty errors were made during forward
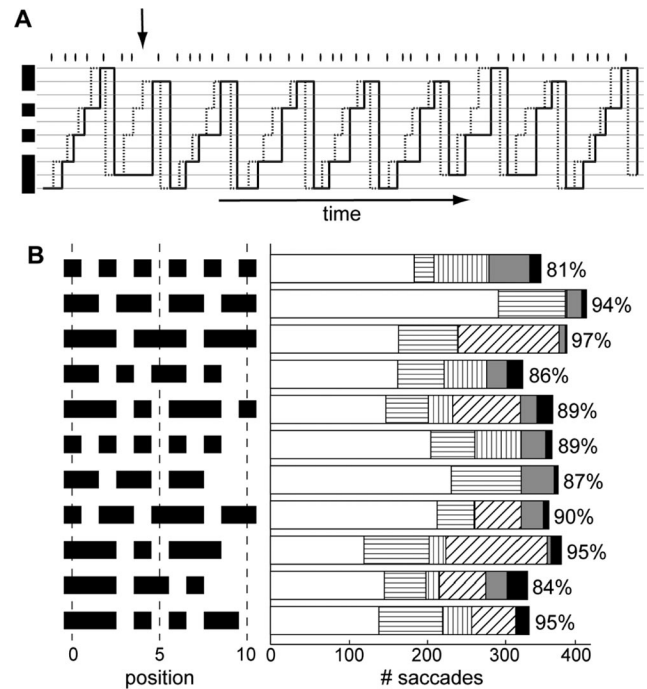

Figure 9. Error statistics. (A) Example of a single trace of the position of the attention focus (dotted line) and gaze (solid line) versus time over a period of 10 s. The network reads the line shown on the left over the whole period. Word recognitions are indicated by the marks above. The arrow indicates an erroneous shift of attention that was not preceded by recognition of a word. (B) Statistics of reading for several lines. Each bar summarizes the saccades made during 100 s of reading of the line of text shown on the left (white: correct forward saccades; horizontal lines: correct return saccades; vertical lines: correct skippings of a word; oblique lines: double fixations on long words; gray: errors in right-to-left saccades; black: errors in forward saccades). Numbers on the right give the percentage of correct saccades.

saccades, and of those, 16 were due to double recognitions of words. Saccades starting from the second-to-last word were particularly prone to this error when the network tried to skip the last word, but could not find the next target, as no words remained within the reading bias. Two errors were skippings of a short word with unclear recognition of that word, and the remaining two instances were classified as errors because of fluctuations in Layer 2/3. In these cases, attention was not clearly allocated.

In the seventh line (*xx xx xx*) of Figure 9B, 13% (47/366) of the saccades were errors. Remarkably, most of them (44/47) were made in saccades to the left (mainly erroneous return saccades, but also some other saccades to the left). These errors illustrate one shortcoming of the current model. Forty of them occurred on the first letter of the last word in the line. At this position, not a single letter of the string was within the bias from Layer 6r to Layer 4 (cf. Figure 3A), because the bias for return saccades targeted only Positions -7 to -10, even further left than the beginning of the line. Hence, the network randomly selected the first or the second word as the next target. The remaining errors (left to right and right to left) were due to erroneous selections and fluctuations in the attentional signal.

Visual Scanning Without the Reading Rule

The global behavior of the network (i.e., whether it should read or simply scan) was defined by the rule input that targeted Layer 6r of the FEF. When this input was inactive, the bias for left-to-right eye movements disappeared, and the network scanned the one-dimensional visual scene (see Figure 10A). The network was simulated for 1,000 s with the same visual input array of six targets with relative strengths (1, 0.9, 0.8, 0.9, 1, 0.8). Cognitive (word) processing was restricted to the fovea in these simulations and occurred with the same speed as for long words. However, the speed of word processing did not influence the fixation pattern but only changed single fixation times. Average fixation time for this setting was 325 ± 89 ms, which again roughly corresponds to the observed fixation times in human visual search (Henderson, 2003).

The spatial distribution of fixations reflected the bottom-up saliency of the visual input, as given by the strength of dorsal stream visual input to Layer 4, without any top-down bias (see Figure 10B). The model exhibited an inhibition of return that assured that the network was not simply alternating gaze between the two most salient stimuli. Inhibition was achieved by a bias in the computation of visual saliency (see Appendix, Connection [5]; for an in-depth description of the mechanism of inhibition of return, see Heinzle et al., 2007).

A second interesting test that demonstrated the influence of the rule input was when the visual input was taken away from the network. Under this condition the behavior of the network was driven by spontaneous fluctuations in the network activity that also led to saccadic eye movements. With reading rule input, saccades were rather frequent (fixation times: 533 ± 288 ms) and closely followed the spatial bias (see Figure 11A). Most saccades were initiated by fluctuations in Layer 4. Without rule input, most of the

effective fluctuations originated in Layer 2/3. Fixation times were much longer ($4,285 \pm 3,932$ ms), and the saccade vector distribution was flat (Figure 11B), showing again that the network was not biased when the rule input was turned off.

Discussion

A biologically realistic cortical model was used to control sequential saccadic eye movements during reading. The dynamics within the model relied on the interaction of spiking neurons. A global rule input that signaled either "read" or "scan" could change the network's behavior dramatically, but for both tasks the network produced realistic distributions of fixation times. In the case of reading, interaction with a cortical WP module that processed words allowed the network to read an array of words efficiently, and to skip short words, as human readers do. Word processing and saccade buildup were both modeled by a race to threshold. The results show that it is indeed possible to tightly couple processes of oculomotor control and cognition in a realistic network of spiking neurons.

In the implementation of the reading mechanisms, we took great care to maintain the basic circuit of the FEF as detailed in Heinzle et al. (2007). In that instantiation, the model was assumed to be "alexic," and instead was instructed to perform saccadic tasks to cues that were laid down in a command recognition module. Here the model was extended by adding a WP module, which gave it the competence to read. It is important to note that now the very same cortical circuit controlled the oculomotor behavior in two quite different tasks: reading and scanning. This demonstrates the computational flexibility of the canonical cortical circuit and shows that, in principle, few modifications suffice to adapt a given cortical circuit for seemingly very different tasks. It has been suggested that successful "cultural recycling" of given cortical structures is the key for recently (within the last few thousand years) developed human cognitive functions such as arithmetic and reading (Dehaene & Cohen, 2007). Our model suggests that the existing structure of the FEF together with cultural recycled ventral areas that process words can perform reading eye movements. Indeed, only minor changes are required to adapt the behavior of the model FEF for monkeys to the reading model FEF presented here. In fact, it would have been worrying if the FEF model proposed for the monkey was not suitable for the human task of reading, as monkeys' and humans' visual systems are very similar, and reading is a relatively recent addition to the repertoire of human behavior. Although this does not prove any universality of the canonical circuit, it at least demonstrates how similar cortical circuits can be adapted to compute new functions. Our model, which couples just two cortical areas, is highly simplified compared with the actual network involved in fluent reading, which is composed of many cortical and subcortical areas (Salmelin, Helenius, & Service, 2000). The dorsal cortical network involved in the control of saccadic eye movements includes not only the FEF but also the parietal eye fields, the supplementary eye fields, and the visual areas of the dorsal stream (for a review, see Büttner & Büttner-Ennever, 2006). The ventral cortical network for the processing of words also includes several early visual areas and the VWF area, as well as Broca's area and a prerolandic area (Vinckier et al., 2007). Our model captures these two processing streams to demonstrate some key aspects of the control of eye movements
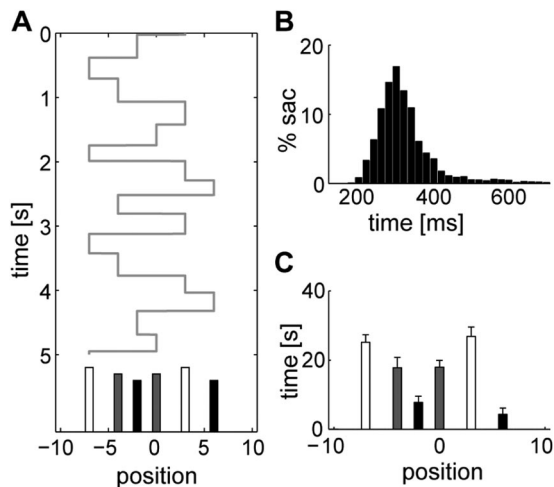
Figure 10. Scanning a visual scene. (A) Sample fixation trace over 5 s. Time runs from top to bottom. Bars below indicate the position and the strength of the visual input of the dorsal cortical stream for the six targets (1, 0.9, 0.8, 0.9, 1, 0.8). The shading of the bars indicates the strength of the input. (B) Fixation time distribution during scanning ($N = 3,069$). (C) Mean fixation time per target. Bars show the mean fixation time per target as the average over 10 simulations of 100 s. Error bars are standard deviations. Shading of the bars indicates the strength of the visual input (cf. Figure 10A). Sac = saccade.

Figure 11. Network simulation without any visual input. (A) Saccade (sac) vector (left) and fixation (fix) time (right) distributions with reading rule input. Average fixation time: 533 ± 288 ms. (B) Saccade vector (left) and fixation time (right) distributions without reading rule input. Average fixation time: 4,285 ± 3,932 ms.

during reading. Although the dorsal stream and in particular the FEF are important for guiding eye movements, the ventral stream (represented by the WP module in our model) comprises areas that process shape.

Reading Performance of the Network
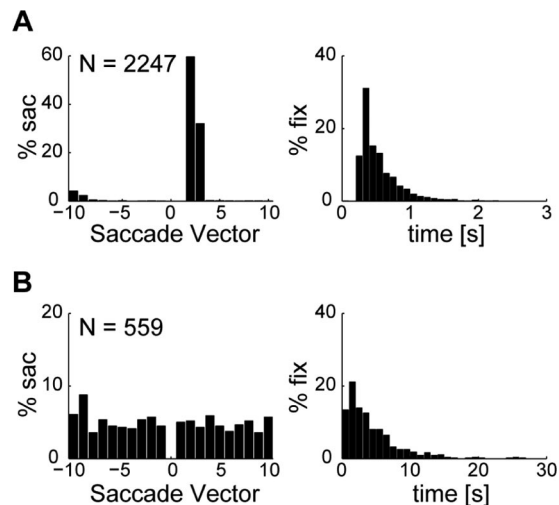
Partly because the network was biologically realistic and constrained by the known cortical architecture (Binzegger et al., 2004; Douglas & Martin, 2004), there were some limitations to it, which need consideration here. Clearly, the spatial resolution of only 21 horizontally aligned populations limited the size and length of possible input strings. A consequence of this was the highly simplified "language" model that was used. A real lexical processing module would probably allow the network to capture more details of real eye movement traces, such as word frequency effects, and could also be used to introduce regressive saccades (see, e.g., Rayner, 1998). A network of canonical cortical circuits for human linguistic processing (Jackendoff, 2007) is, however, still far away. Hence, word processing in the WP module of the model relied on word length solely. Nevertheless, in common with all other models (for an overview, see Reichle et al., 2003), the FEF model reproduced well the general fixation time statistics for reading.

The retinotopic mapping of the model was implemented by a uniform spacing between different positions. This does not capture the segregation of the FEF into small and large saccade regions (Schall et al., 1995). Large return-sweep saccades are most probably guided by the large saccade region, and normal reading saccades by the small saccade region. Because our model did not incorporate such a division of the FEF, the bias of visual selection for forward saccades relied on the same mechanism as the bias for return saccades. A significant part of the errors of the network, in particular in return saccades, was a direct consequence of the small differences in bias between different retinotopic positions.

Presently, it is not possible to compare directly the detailed error classification of the model's behavior with human data. Nevertheless, we checked for errors directly in the combined pattern of eye movements, attention, and word processing. We do not know of any study that could directly measure either visual attention or the recognition of words in a normal reading task.[4] The only measures to assess errors in human experiments are content-related questions. So the performance measure, which was generally between 85% and 90% correct saccade patterns, simply reflected how well the network reproduced the ideal sequence of attention, recognition of words, and eye movements, as defined in Figure 1. Given that the network was running in a noisy environment, as also demonstrated by the simulations without any visual input, this is a good performance, as it would be easy to make the model more deterministic by reducing the internal noise. The model is in general very robust to small changes in strength of individual connections. A detailed discussion of the robustness of the FEF model is beyond the scope of this article but can be found in Heinzle (2006).

Comparison to More Functional Models of Reading Eye Movements

One of the major goals of our model was to demonstrate how functional models of eye movement control during reading (for an overview, see Reichle et al., 2003) could be implemented in a biologically realistic structure. We focus here on the serial-processing E-Z Reader model (Reichle et al., 2006) and the more parallel SWIFT model (Engbert et al., 2005; Richter et al., 2006). At this point, it has to be said that the model presented here was not intended to capture every detail of eye movements during reading, but rather it presents a first step toward a biologically realistic implementation of the psychological models. The need for such a kind of validation has been pointed out by the architects of the E-Z Reader model (Reichle et al., 2003), who speculated about a possible neural implementation. It is difficult to compare the two models directly, but we think there are some points that should be mentioned. Both models rely on a tight interaction between the dorsal and the ventral stream of visual processing. However, where this interaction happens differs between the two. The architects of the E-Z Reader model suggest that it occurs at the stage of the parietal eye fields, whereas the FEF model suggests that it is the well-documented connection from the FEF to the temporal lobe (Moore & Armstrong, 2003; Moore & Fallah, 2004; Schall et al., 1995) that is crucial. Although both the parietal cortex and the FEF are associated with attention (Corbetta & Shulman, 2002), the FEF seems to be linked more directly to eye movements and shifts of covert attention. These attentional and eye movement signals recorded in the FEF make it, in our view, the prime candidate region to mediate this tight interaction.

Our model implemented serial processing in the selection of attention (Morrison, 1984; Reichle et al., 1998). Such attentional shifts were necessary for foveation (Deubel & Schneider, 1996;

---

[4] Gaze-contingent eye movement experiments can be used to measure the span of attention and to infer whether words have been recognized (for a review, see Rayner, 2009). Nevertheless, they provide only an indirect measure of the two.

Kowler et al., 1995). The visual saliency was active on a large span, but we did not implement a transsaccadic remapping of salient targets (Balan & Ferrera, 2003; Sommer & Wurtz, 2006; Umeno & Goldberg, 1997). To be able to build a WP module with the same neuronal elements as for the FEF model, word processing was simplified to rely solely on word length (Rayner & McConkie, 1976). Obviously, this implementation of word processing does not capture the influence of word frequency (Schilling, Rayner, & Chumbley, 1998) and information content (Rayner, 1977) on fixation times in reading (for reviews, see Rayner, 1998, 2009). Nevertheless, the use of word length as an approximation for the lexical difficulty of a word seems to be reasonable. For word skipping, human experiments suggest that word length is the dominating factor (Brysbaert, Drieghe, & Vitu, 2005), although how predictable a word is from prior context also influences skipping (Rayner, 1998). It is important to note that the average recognition time, which was defined as the time span between the first allocation of attention until the time when the WP module signaled that attention could be withdrawn from the word, had to be rather short (in the range of 150 ms). Hence, word recognition in our model has to be interpreted in a similar way as in the first stage of processing in the E-Z Reader model (Reichle et al., 1998, 2003). Clearly, from the simplified language and word recognition mechanism, we cannot draw any conclusions about the timing of lexical processing.

Parallel processing models (Engbert et al., 2005; Richter et al., 2006) that consider the timing of eye movements to be independent of word processing have to rely on a more elaborate computation of visual saliency, one that incorporates a lexical access parameter over a wide perceptual span. It is not straightforward to implement such mechanisms in our model because it is not obvious how to model lexical access in a spiking network. Nevertheless, given a way to calculate the saliency according to lexical access, this saliency information could be integrated into our model in two ways: either directly via the bottom-up input or through the flexible top-down rule input. Finally, in a parallel processing version of the model there should not be any direct coupling in timing between the FEF and the WP module, so that there is no need for a word recognition signal to initiate a shift of attention.

There is one important timing difference between our model and the experimental data. Prior to the skipping of a word, fixation times were increased by approximately 180 ms in our model. Experiments show that this effect is around 30 ms (Pollatsek et al., 1986; Rayner et al., 2004) in normal reading. Benefits and costs of skipping have been important in comparing the two main models for eye movements in reading. Both E-Z Reader and SWIFT show a general increase in fixation time before skippings (Engbert et al., 2005; Pollatsek et al., 2006). A detailed statistical analysis of a large corpus showed that fixation times prior to skippings depend on the length and frequency of the skipped word (Kliegl & Engbert, 2005). The SWIFT model captures this word length effect (Engbert et al., 2005), but it also fails to model a reduction in fixation time prior to the skipping of short words. Our model shows an increase prior to skippings with far too high of a skipping cost. This is due to the reprogramming of the saccade including the second shift of attention. Although the model captured well the overall fixation time distribution in reading, this increase in fixation time prior to skippings suggests that some additional mecha-

nism is needed to speed up saccade planning after serial shifts of attention. One means of implementing such a speedup would be a release of inhibition in the motor buildup Layer 5r. A global inhibition that decreases over time after each saccade would result in much faster saccadic buildup for the saccades following a second shift of attention, and therefore would reduce the inflation of fixation times prior to skipping. A second possible implementation could be to allow several positions in Layer 5r to build up activity simultaneously, so that if one saccade plan gets canceled, the neural population of a different position would already fire at an elevated rate and therefore take less time to trigger a saccade. It is interesting that the FEF model predicted quite accurately the increase in fixation time after skipping, which is of the order of 30–50 ms in experiments (Kliegl & Engbert, 2005; Rayner, 2009).

Reading requires a tight interaction of the allocation of attention, eye movements, and cognitive processes. Our model captured such an interaction by a competition in a race-to-threshold mechanism (Logan & Cowan, 1984). A similar mechanism for oculomotor control was used successfully to reproduce fixation times in reading (Carpenter & McDonald, 2007; McDonald et al., 2005).

Attention always preceded the eye movements, and therefore the model was in line with a premotor theory of attention (Deubel & Schneider, 1996; Kowler et al., 1995; Rizzolatti, Riggio, Dascola, & Umiltà, 1987). Similar attentional effects have been recorded in humans during sequential eye movements (Gersch, Kowler, & Dosher, 2004). It is important, however, that attention and eye movements were controlled independently, as in the skipping of a word. Such an independent control has been shown for the FEF in monkeys (Thompson, Biscoe, & Sato, 2005) and was inherent in our model.

A crucial requirement in the guidance of eye movements in cognitively demanding tasks is the coordination of position and time, the where and the when of the saccade (Findlay & Walker, 1999). The FEF model implements this interaction at the level of attention. While the dorsal path together with the "reading bias' defined the spatial position to be attended, the ventral path provided a timing signal for shifting attention. In addition, activation of the fixation neurons during foveal attention prevented the motor output (Layer 5) from building up any activity as long as attention was required at the fovea. All these interactions happened at the stage of the FEF, whereas in the model proposed by Findlay and Walker (1999), the crucial interaction between the where and the when occurs in the superior colliculus.

Model Predictions

Because of the simplified view of word recognition, we do not think that the current model can make new experimental predictions at the level of reading eye movements. However, the detailed modeling allows us to make some predictions on the level of neurons and neural populations. It is clear that methods presently available for noninvasive research in humans are not sufficient. Either they are too slow (functional MRI [fMRI]) or they have a too low spatial resolution (electroencephalography [EEG] or magnetoencephalography) to examine the details of cortical microcircuit activations in reading. However, in the future, higher spatial resolution in fMRI (Yacoub, Harel, & Ugurbil, 2008), the combination of fMRI and EEG (Laufs & Duncan, 2007), electrocorticograms (Pistohl, Ball, Schulze-Bonhage, Aertsen, & Mehring,

2008) and single-cell recordings (Quian Quiroga, Kreiman, Koch, & Fried, 2007; Sahin, Pinker, Cash, Schomer, & Halgren, 2009) from patients may provide the tools to measure some key aspects of human neural responses that are relevant for our model.

Hence, the model can make some predictions for future physiological experiments. The tight coupling between activity related to attentional shifts (in the FEF) and recognition of words (in some temporal cortical word processing region, such as the VWF region) should be measurable, particularly when words are skipped. It might also be possible to separate the race-to-threshold behavior in the WP module and the saccade buildup. Some attempts have been made to reveal the cognitive processes during reading using EEG (Sereno & Rayner, 2003). Fast motor buildup before saccades was observed with EEG in the human eye fields (Yamamoto et al., 2004). Recent fMRI experiments demonstrated the buildup of activity in artificially slowed-down recognition (Ploran et al., 2007) and antisaccade tasks (Medendorp, Goltz, Crawford, & Vilis, 2005).

A further prediction arises from the top-down bias from Layer 6r to Layer 4 in the FEF. This bias is learned in the anatomical connections of the model. Therefore, it should be reflected in other tasks than reading, as long as the reading bias is somehow active. The bias is observed in a task where left-to-right motion is involved (Spalek & Hammad, 2005). In addition, the fact that the distribution of saccade lengths is the same for single and repeated reading of text (Schnitzer & Kowler, 2006) also points in that direction. Finally, the simulation of the model without visual input suggests that people who just imagine reading should also produce a similar eye movement pattern. Some studies examined imagined, so-called mindless reading by replacing all letters in a sentence by a *z* (Nuthmann & Engbert, 2009; Rayner & Fischer, 1996; Vitu, O'Regan, Inhoff, & Topolski, 1995). A slight increase in fixation times was observed. However, subjects could rely on the still-visible spatial information of the words. In our suggested experiment, subjects should imagine reading without any visual input or, if needed, on a line where blank spaces are also replaced by a *z*. The reading bias of the visual saliency should also be reflected in the activation of the FEF, and therefore it should also be visible in the topographic activity of the FEF.

An alternative to testing the model with humans in real reading might be to train monkeys in tasks that are directly comparable to reading and to record from neurons in the FEF. Monkeys, for example, could be trained to scan an array of oriented gratings until they find a specific target orientation. Recognition speed of individual gratings could be manipulated by changing the spatial frequency or contrast of the grating and therefore allowing the monkey to process it in the parafovea or not.[5] Other reading-like tasks have already been studied in humans (Trukenbrod & Engbert, 2007) and could be adapted for monkeys.

General Conclusions

Although Reichle et al (2003) speculated about a possible neural implementation of the E-Z Reader model, the network model presented here is to our knowledge the first successful attempt to bridge the gap between detailed, biologically realistic models of neural networks and psychological models of reading, and still reproduce to a large extent the measured statistics of reading eye movements (Rayner, 1998, 2009). Because any cortical computa-

tion has to rely on the hardware of the neocortex, it is an important question how high-level models of cognitive functions can be implemented by biological neural networks that respect the anatomy of the layered cortical architecture and include spiking neurons as well as realistic synapses. We think that the FEF model provides us with a first step that shows how high-level models can be based directly on cortical circuits of spiking neurons. The layered canonical circuit, inferred from the anatomy of cat visual cortex (Binzegger et al., 2004), could be conserved and successfully guide eye movements in reading, as shown here. But it was also able to control a variety of behaviors when tested in the context of the monkey FEF (Heinzle et al., 2007). Arguably, within the next decades, experimental methods will allow us to test directly the structure and dynamics of this model in reading experiments in humans. Detailed computational models will help in understanding neocortical structure and function by linking higher cognitive functions to basic cortical circuits.

---

[5] See also the discussion in Heinzle (2006) for the description of potential reading-like tasks for monkeys.

References

Abeles, M. (1991). *Corticonics: Quantitative study of the cerebral cortex.* New York, NY: Cambridge University Press.

Armstrong, K. M., Fitzgerald, J. K., & Moore, T. (2006). Changes in visual receptive fields with microstimulation of frontal cortex. *Neuron, 50*(5), 791–798.

Balan, P. F., & Ferrera, V. P. (2003). Effects of gaze shifts on maintenance of spatial memory in macaque frontal eye field. *Journal of Neuroscience, 23*(13), 5446–5454.

Barone, P., Batardiere, A., Knoblauch, K., & Kennedy, H. (2000). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *Journal of Neuroscience, 20*(9), 3263–3281.

Binzegger, T., Douglas, R. J., & Martin, K. A. C. (2004). A quantitative map of the circuit of cat primary visual cortex. *Journal of Neuroscience, 24*(39), 8441–8453.

Braitenberg, V., & Schüz, A. (1991). *Anatomy of the cortex: Statistics and geometry.* Berlin, Germany: Springer-Verlag.

Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 53–77). Oxford, England: Oxford University Press.

Büttner, U., & Büttner-Ennever, J. A. (2006). Present concepts of oculomotor organization. In J. A. Büttner-Ennever (Ed.), *Progress in Brain Research: Vol. 151. Neuroanatomy of the Oculomotor System* (pp. 1–42): Amsterdam, the Netherlands: Elsevier.

Carpenter, R. H., & McDonald, S. A. (2007). LATER predicts saccade latency distributions in reading. *Experimental Brain Research, 177*(2), 176–183.

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *NeuroImage, 40*(1), 353–366.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201–215.

Dehaene, S. (2009). *Reading in the brain: The science and evolution of human invention.* New York, NY: Viking.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384–398.

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

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*(12), 1827–1837.

Douglas, R. J., & Martin, K. A. C. (2004). Neuronal circuits of the neocortex. *Annual Review of Neuroscience, 27*, 419–451.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42*(5), 621–636.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777–813.

Everling, S., & DeSouza, J. F. X. (2005). Rule-dependent activity for prosaccades and antisaccades in the primate prefrontal cortex. *Journal of Cognitive Neuroscience, 17*(9), 1483–1496.

Feng, G. (2006). Eye movements as time-series random variables: A stochastic model of eye movement control in reading. *Cognitive Systems Research, 7*(1), 70–95.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences, 22*(4), 661–721.

Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., . . . Cohen, L. (2006). Direct intracranial, fMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron, 50*(2), 191–204.

Gersch, T. M., Kowler, E., & Dosher, B. (2004). Dynamic allocation of visual attention during the execution of sequences of saccades. *Vision Research, 44*(12), 1469–1483.

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

Heinzle, J. (2006). *A model of the local cortical circuit of the frontal eye fields* (Doctoral dissertation). Swiss Federal Institute of Technology, Zurich, Switzerland. (ETH Diss. No. 16897)

Heinzle, J., Hepp, K., & Martin, K. A. C. (2007). A microcircuit model of the frontal eye fields. *Journal of Neuroscience, 27*(35), 9341–9353.

Heller, D. (1982). Eye movements in reading. In R. Groner & P. Fraisse (Eds.), *Cognition and eye movements* (pp. 139–154). Amsterdam, the Netherlands: North-Holland.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*(11), 498–504.

Henderson, J. M., & 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.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution.* Oxford, England: Oxford University Press.

Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain Research, 1146*, 2–22.

James, K. H., & Gauthier, I. (2006). Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia, 44*(14), 2937–2949.

Kisvarday, Z. F., Cowey, A., Smith, A. D., & Somogyi, P. (1989). Interlaminar and lateral excitatory amino acid connections in the striate cortex of monkey. *Journal of Neuroscience, 9*(2), 667–682.

Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychonomic Bulletin & Review, 12*(1), 132–138.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35*(13), 1897–1916.

Kritzer, M. F., & Goldman-Rakic, P. S. (1995). Intrinsic circuit organization of the major layers and sublayers of the dorsolateral prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology, 359*(1), 131–143.

Kujala, J., Pammer, K., Cornelissen, P., Roebroeck, A., Formisano, E., & Salmelin, R. (2007). Phase coupling in a cerebro-cerebellar network at 8–13 Hz during reading. *Cerebral Cortex, 17*(6), 1476–1485.

Laufs, H., & Duncan, J. S. (2007). Electroencephalography/functional MRI in human epilepsy: What it currently can and cannot do. *Current Opinion in Neurology, 20*(4), 417–423.

Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr. Chips: An ideal-observer model of reading. *Psychological Review, 104*(3), 524–553.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*(3), 295–327.

McDonald, S. A., Carpenter, R. H., & Shillcock, R. C. (2005). An anatomically constrained, stochastic model of eye movement control in reading. *Psychological Review, 112*(4), 814–840.

Medendorp, W. P., Goltz, H. C., Crawford, J. D., & Vilis, T. (2005). Integration of target and effector information in human posterior parietal cortex for the planning of action. *Journal of Neurophysiology, 93*(2), 954–962.

Moore, T., & Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature, 421*(6921), 370–373.

Moore, T., & Fallah, M. (2004). Microstimulation of the frontal eye field and its effects on covert spatial attention. *Journal of Neurophysiology, 91*(1), 152–162.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 667–682.

Neggers, S. F., Huijbers, W., Vrijlandt, C. M., Vlaskamp, B. N., Schutter, D. J., & Kenemans, J. L. (2007). TMS pulses on the frontal eye fields break coupling between visuospatial attention and eye movements. *Journal of Neurophysiology, 98*(5), 2765–2778.

Nuthmann, A., & Engbert, R. (2009). Mindless reading revisited: An analysis based on the SWIFT model of eye-movement control. *Vision Research, 49*(3), 322–336.

Pistohl, T., Ball, T., Schulze-Bonhage, A., Aertsen, A., & Mehring, C. (2008). Prediction of arm movement trajectories from ECoG-recordings in humans. *Journal of Neuroscience Methods, 167*(1), 105–114.

Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., & Wheeler, M. E. (2007). Evidence accumulation and the moment of recognition: Dissociating perceptual recognition processes using fMRI. *Journal of Neuroscience, 27*(44), 11912–11924.

Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inference about eye movement control from the perceptual span in reading. *Perception & Psychophysics, 40*(2), 123–130.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*(1), 1–56.

Pouget, P., Emeric, E. E., Stuphorn, V., Reis, K., & Schall, J. D. (2005). Chronometry of visual responses in frontal eye field, supplementary eye field, and anterior cingulate cortex. *Journal of Neurophysiology, 94*(3), 2086–2092.

Pouget, P., Stepniewska, I., Crowder, E. A., Leslie, M. W., Emeric, E. E., Nelson, M. J., & Schall, J. D. (2009). Visual and motor connectivity and the distribution of calcium-binding proteins in macaque frontal eye field: Implications for saccade target selection. *Frontiers in Neuroanatomy, 3*(Article 2). doi:10.3389/neuro.05.002.2009

Quian Quiroga, R., Kreiman, G., Koch, C., & Fried, I. (2007). Sparse but not "grandmother-cell" coding in the medial temporal lobe. *Trends in Cognitive Sciences, 12*(3), 87–91.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition, 5*(4), 443–448.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 371–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 720–732.

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 58*(5), 734–747.

Rayner, K., Juhasz, B. J., & Brown, S. J. (2007). Do readers obtain preview benefit from word $n + 2$? A test of serial attention shift versus distributed lexical processing models of eye movement control in reading. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 230–245.

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research, 16*(8), 829–837.

Reichle, E. D., & Laurent, P. A. (2006). Using reinforcement learning to understand the emergence of "intelligent" eye-movement behavior during reading. *Psychological Review, 113*(2), 390–408.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125–157.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research, 7*(1), 4–22.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparison to other models. *Behavioral and Brain Sciences, 26*(4), 445–526.

Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research, 7*(1), 34–55.

Richter, E. M., Engbert, R., & Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems Research, 7*(1), 23–33.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2*(11), 1019–1025.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltà, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia, 25*(1A), 31–40.

Ruff, C. C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Freeman, E., Haynes, J.-D., . . . Driver, J. (2006). Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. *Current Biology, 16*(15), 1479–1488.

Sahin, N. T., Pinker, S., Cash, S. S., Schomer, D., & Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science, 326*(5951), 445–449.

Salinas, E. (2003). Background synaptic activity as a switch between dynamical states in a network. *Neural Computation, 15*(7), 1439–1475.

Salmelin, R., Helenius, P., & Service, E. (2000). Neurophysiology of fluent and impaired reading: A magnetoencephalographic approach. *Journal of Clinical Neurophysiology, 17*(2), 163–174.

Sato, T. R., & Schall, J. D. (2003). Effects of stimulus–response compatibility on neural selection in frontal eye field. *Neuron, 38*(4), 637–648.

Schall, J., Morel, A., King, D. J., & Bullier, J. (1995). Topography of visual cortex connections with frontal eye field in macaque: Convergence and segregation of processing streams. *Journal of Neuroscience, 15*(6), 4464–4487.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26*(6), 1270–1281.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G.. (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology, 79*(6), 3272–3278.

Schnitzer, B. S., & Kowler, E. (2006). Eye movements during multiple readings of the same text. *Vision Research, 46*(10), 1611–1632.

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences, 7*(11), 489–493.

Silvanto, J., Lavie, N., & Walsh, V. (2006). Stimulation of the human frontal eye fields modulates sensitivity of extrastriate visual cortex. *Journal of Neurophysiology, 96*(2), 941–945.

Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature, 444*(7117), 374–377.

Spalek, T. M., & Hammad, S. (2005). The left-to-right bias in inhibition of return is due to the direction of reading. *Psychological Science, 16*(1), 15–18.

Taylor, P. C., Nobre, A. C., & Rushworth, M. F. (2007). FEF TMS affects visual cortical activity. *Cerebral Cortex, 17*(2), 391–399.

Thompson, K. G., Biscoe, K. L., & Sato, T. R. (2005). Neuronal basis of covert spatial attention in the frontal eye field. *Journal of Neuroscience, 25*(41), 9479–9487.

Trukenbrod, H. A., & Engbert, R. (2007). Oculomotor control in a sequential search task. *Vision Research, 47*(18), 2426–2443.

Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *Journal of Neurophysiology, 78*(3), 1373–1383.

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron, 55*(1), 143–156.

Vitu, F., O'Regan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless reading: Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception & Psychophysics, 57*(3), 352–364.

Yacoub, E., Harel, N., & Ugurbil, K. (2008). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences, USA, 105*(30), 10607–10612.

Yamamoto, J., Ikeda, A., Satow, T., Matsuhashi, M., Baba, K., Yamane, F., . . . Shibasaski, H. (2004). Human eye fields in the frontal lobe as studied by epicortical recording of movement-related cortical potentials. *Brain, 127*(4), 873–887.

Yang, S. (2006). An oculomotor-based model of eye movements in reading: The competition/interaction model. *Cognitive Systems Research, 7*(1), 56–69.

(Appendix follows)

Appendix

Detailed Mathematical Description of the Model

This appendix provides a detailed description of the model neurons and the connections within the network. The dynamics of the network was given by the dynamics of the single neurons and their interactions through conductance-based synapses. First, the individual elements (neurons and synapses) are described, and then the connectivity pattern of the model is given in detail. It is important to note that the model presented here builds on the same microcircuit that guided single saccades (Heinzle et al., 2007). All adaptations that were made to implement reading leave intact the functionality of the original frontal eye field (FEF) model. The rapid switching between different behaviors depends entirely on the gating of functional connections due to external rule inputs and does not rely on any rewiring or synaptic changes.

General Network Architecture

The architecture of the FEF local circuit model was built on some general features of cortical architecture. The relative proportion of excitatory and inhibitory neurons reflected the 4:1 ratio observed in cortex (Abeles, 1991; Braitenberg & Schüz, 1991). The synaptic connections within the network followed two main anatomical principles. First, strong intralaminar connections and weak feed-forward interlaminar connections were inserted as suggested by experimental data (Binzegger et al., 2004; Kisvarday, Cowey, Smith, & Somogyi, 1989; Kritzer & Goldman-Rakic, 1995). Second, this network was tuned to scan an array of targets and produce single saccades comparable to those in primate behavior. This manual tuning required the insertion of some additional connections.

Figure A1 shows the structure of the network (cf. Figure 2). Each layer of the FEF circuit contained several populations of integrate-and-fire (IF) neurons located at 21 retinotopic positions along the horizontal axis. Each retinotopic position in Layer 4 and in Layer 2/3 contained a population of 100 excitatory and 25 inhibitory neurons. In Layer 5, populations of 40 excitatory and 25 inhibitory ramping neurons (Layer 5r) and the same numbers of bursting neurons (Layer 5b) were inserted.

Layer 6 consisted of excitatory populations of 50 attention-related (6r) and 50 saccade-related (6s) neurons at each retinotopic position. A single foveal population of inhibitory neurons was inserted in Layer 6. Recurrent inhibition in Layer 6 allowed the foveal population within Layer 6 to be active whenever attention was on the fovea or no attention was allocated at all. Finally, one population of 100 fixation neurons was included in the network.

Neurons and Synapses

The basic elements of the FEF model, IF neurons and synapses, were defined similar to those of Salinas (2003). The membrane dynamics of the IF neurons were given by

$$\tau_m \frac{dV(t)}{dt} = -V(t) + g_e(t)(V(t) - V_e) + g_i(t)(V(t) - V_i).$$

(1)

Crossing of a threshold V_{th} triggered a spike followed by a reset of the membrane voltage to V_r and a refractory period of t_r. The single-cell parameters are given in Table A1.

The conductances g_e and g_i consisted of two parts. First, synapses within the FEF were modeled as decaying exponential conductances:

$$\tau_{e,i} \frac{dg_{e,i}}{dt} = -g_{e,i}.$$

(2)

Each spike instantaneously increased the conductance of the corresponding synapse by a fixed weight: $g_{e,i} \rightarrow g_{e,i} + w_{e,i}$. Weights and time constants differed between connections and are described in a separate paragraph (see also Table A3).

Second, the external inputs to each neuron were modeled as fluctuating conductances $g_{ext}(t)$ (Salinas, 2003) and added to the internal conductances:

$$\tau_{ext} \frac{dg_{ext}}{dt} = -(g_{ext} - \mu_{e,i}) + \sqrt{D}\chi(t).$$

(3)

The fluctuations of the external input were given by the diffusion constant $D = (\mu_{e,i}w_{e,i}/\tau_{ext})^{1/2}$ and a white Gaussian noise $\chi(t)$; $\mu_{e,i}$ was the mean conductance of the external input. The external weights $w_e = 0.02$ and $w_i = 0.06$ and the time constant $\tau_{ext} = 3$ ms defined the size and the temporal correlation of the input. Background inputs drove single neurons to spontaneous firing rates of less than 10 Hz. Only the fixation neurons had a spontaneous firing rate of around 40 Hz. The μ values of all the external inputs are given in Table A2.

The dorsal visual input to Layer 4 was turned on 50 ms (Schmolesky et al., 1998) after presentation of the stimulus on the screen, or after the last saccade, and it was reduced in intensity to 50% of the initial value 40 ms later until the next saccade. This temporal pattern roughly captured the transient and sustained responses to visual stimulation. The spatial pattern of the visual input was given by the relative strength of the inputs at each retinotopic position. If the population activity of bursting neurons in Layer 5b crossed a threshold of 50 Hz, a saccade to the corresponding retinotopic position was initiated, and the visual input was updated accordingly. The fixation input targeted the population of fixation neurons. It was turned off 50 ms after the offset of the fixation stimulus on the screen. During reading and scanning, there was no external fixation input given to the network.
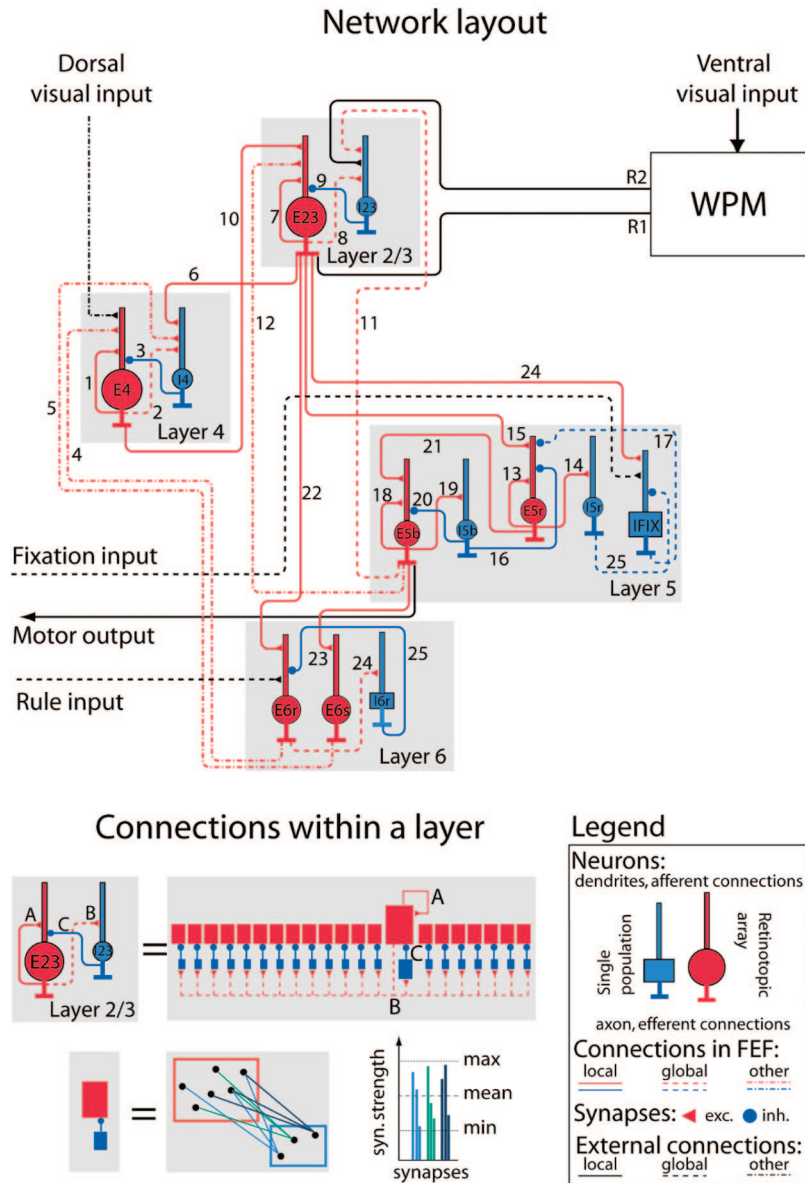
Figure A1. Top: Complete layout of the architecture of the frontal eye field (FEF) model circuit. Colored circles are full retinotopic representations consisting of arrays of 21 populations of neurons. Colored rectangles are single populations (e.g., fixation neurons; red neurons are excitatory [exc.], and blue neurons are inhibitory [inh.]). External inputs and outputs of the FEF are shown in black. Layer 4 received a dorsal visual input, which was not feature specific. The input to the fixation neurons was not used during reading. The rule input to Layer 6r switched the network behavior between reading and visual search. Bursting neurons in Layer 5b provided the motor output of the FEF. The spatial pattern of the connections is summarized into three groups. *Local* connections (solid lines) connected only to populations at the same retinotopic position, whereas *global* connections (dashed lines) connected to all retinotopic positions. The connections that could not be grouped into one of the two above were called *other* connections (dash-dotted line). A detailed list of all connections within the FEF is given in Table A3. The connections to and from the word processing module (WPM) are shown in black. The WPM received a feature-specific visual input, which represented the ventral processing stream. Layer 2/3 connected to the WPM [R1] and in turn received input from it [R2]. Bottom: Retinotopic arrangement of the connections within Layer 2/3. (A) Local self-excitation (only shown for one sample retinotopic position). (B) Global excitation of all inhibitory populations (only shown for one efferent excitatory population). (C) Local inhibitory connections. Below, the random selection of connections is illustrated for three sample inhibitory neurons connecting randomly to 50% of the excitatory neurons. The distribution of the weights is indicated by the histogram on the right. The minimum (min), maximum (max), and mean of the uniform distribution are shown by the horizontal dashed lines. Syn. = synapse.

(Appendix continues)

Table A1
Single Neuron Parameters

Neuron	Excitatory	Inhibitory
τ_m	20 ms	10 ms
V_e	74 mV	74 mV
V_i	-10 mV	-10 mV
V_{th}	20 mV	20 mV
V_r	10 mV	10 mV
t_r	1.8 ms	1.2 ms
g_m	25 nS	20 nS

The rule input to Layer 6r targeted all retinotopic positions. It enhanced the firing of Layer 6r cells. In particular, the combination of external rule input and attentional input from Layer 2/3 efficiently drove Layer 6r neurons. Note that the sum of the rule input and the background input to Layer 6r corresponded to the background input of most other populations. The rule input was an external input to the network.

Detailed Connectivity of the FEF

The connections within the network belonged to three main classes: (a) local connections that connected within a single retinotopic position or at most to the nearest neighbors, (b) global connections that connected from each population of one kind of neurons to all populations of another one, and (c) other connections that had a more specific connectivity pattern. Figure A1 shows the network with all its connections. Connections are numbered according to Table A3, and the same numbers in brackets are used to refer to particular connections in this appendix (e.g., [1] is the excitatory connection within Layer 4).

The connection between two classes of cells (e.g., excitatory neurons and inhibitory neurons in Layer 2/3) was described by $W_{nm,pq}^{I23E23}$ (the weight of the synapse from neuron m in the excitatory population q of Layer 2/3 to neuron n in the inhibitory population p of Layer 2/3). The individual synaptic weights were assigned as follows.

A population weight matrix W_{pq}^{I23E23} defined the average weight of the synapses between population q in Layer 2/3 and population p in Layer 2/3. Individual weights were randomly, uniformly distributed between 0.5 and 1.5 times this average weight. The average weights and time constants of all connections are listed in Table A3. In addition, 50% of the weights were randomly set to zero (see Figure A1, bottom). In the E to I connections within Layers 4 [2] and 2/3 [8], 75% of the weights were set to zero, resulting in 25% connectivity. The weights of all connections and their synaptic time constants are given in Table A3.

Local connections (solid lines in Figure A1) were described by the weight matrix

$$W_{pq}^{AB} = w^{AB}\delta_{pq}, \tag{4}$$

with $\delta_{pq} = 1$ if $p = q$ and 0 otherwise. The self-excitation within Layer 4 [1] included a weak nearest neighbor interaction:

$$W_{pq}^{E4E4} = w^{E4E4}[\delta_{pq} + 0.05(\delta_{p(q-1)} + \delta_{p(q+1)})]. \tag{5}$$

The connection from Layer 2/3 excited inhibitory neurons in Layer 4 locally and includes nearest neighbors [6]:

$$W_{pq}^{I4E23} = w^{I4E23}(\delta_{pq} + \delta_{p(q-1)} + \delta_{p(q+1)}). \tag{6}$$

This connection suppressed visual activity in Layer 4 once the attentional focus was on the target selected by Layer 4. Global connections (dashed lines in Figure A1) targeted all retinotopic positions. These connections were fully described by their weight: $W_{pq}^{AB} = w^{AB}$.

Finally, some connections were more specific than the local and global ones described above. Such connections, for example, were involved in biasing the visual selection in Layer 4 or provided an inhibition of return. The connection from Layer 6s neurons to inhibitory neurons in Layer 4 [5] consisted of two parts: a global fast component with weight w^{I4E6s} that reset the activity in Layer 4 after each saccade and a slow ($\tau = 50$ ms) excitation to inhibitory neurons representing the position opposite in the visual field:

$$W_{IR,pq}^{I4E6s} = w_{IR}^{I4E6s}\delta_{p(2z-q)}. \tag{7}$$

Here z is the position of the fovea relative to the left-most position represented in the network. This connection introduced an inhibition of return in the visual selection process of Layer 4 by biasing the probability of visual selection to be low at the position fixated last.

The connection from Layer 6r to Layer 4 excitatory neurons [4] provided the left-to-right bias for reading (see Figure 2). The connection matrix consisted of two parts. One defined the connections that biased visual selections to the right of the currently attended location.

Table A2
Mean Values of Fluctuating External Inputs

Neuronal population	μ_e	μ_i
Layer 4 exc and Layer 2/3 exc	0.472	0.34
Layer 4 inh and Layer 2/3 inh	0.46	0.40
Layer 5r exc	0.45	0.34
Layer 5r inh	0.42	0.34
Layer 5b exc	0.38	0.30
Layer 5b inh	0.32	0.34
Layer 6r exc	0.20	0.34
Layer 6s exc	0.44	0.34
Layer 6r inh	0.455	0.34
Fixation neurons	0.46	0.12
Visual input to Layer 4	0.056	
Rule input to Layer 6	0.24	
Fixation input	0.20	

Note. exc = excitatory; inh = inhibitory.

Table A3

Average Weights of Synapses Within the Local Circuit Model

Connection	To	From	Type	Weight	τ (ms)	Figure A1
$W_{pq}^{E_4E_4}$	L4 exc	L4 exc	Local	0.016	5	1
$W_{pq}^{I_4E_4}$	L4 inh	L4 exc	Global	0.01	5	2
$W_{pq}^{E_4I_4}$	L4 exc	L4 inh	Local	0.12	3	3
$W_{pq}^{E_4E_{6r}}$	L4 exc	L6r exc	Other	0.0024	5	4
$W_{pq}^{I_4E_{6s}}$	L4 inh	L6s exc	Global	0.008	10	5
$W_{IR,pq}^{I_4E_{6s}}$	L4 inh	L6s exc	Other	0.0016	50	5
$W_{pq}^{I_4E_{23}}$	L4 inh	L2/3 exc	Local	0.0028	5	6
$W_{pq}^{E_{23}E_{23}}$	L2/3 exc	L2/3 exc	Local	0.0096	10	7
$W_{pq}^{I_{23}E_{23}}$	L2/3 inh	L2/3 exc	Global	0.008	5	8
$W_{pq}^{E_{23}I_{23}}$	L2/3 exc	L2/3 inh	Local	0.16	3	9
$W_{pq}^{E_{23}E_4}$	L2/3 exc	L4 exc	Local	0.0032	5	10
$W_{pq}^{I_{23}E_{5b}}$	L2/3 inh	L5b exc	Other	0.04	5	11
$W_{pq}^{E_{23}E_{5b}}$	L2/3 exc	L5b exc	Other	0.017	10	12
$W_{pq}^{E_{5r}E_{5r}}$	L5r exc	L5r exc	Local	0.004	50	13
$W_{pq}^{I_{5r}E_{5r}}$	L5r inh	L5r exc	Local	0.03	5	14
$W_{pq}^{E_{5r}E_{23}}$	L5r exc	L2/3 exc	Local	0.0026	5	15
$W_{pq}^{E_{5r}I_{5b}}$	L5r exc	L5b inh	Local	0.04	10	16
$W_{pq}^{E_{5r}I_{FIX}}$	L5r exc	Fix. inh	Global	0.007	3	17
$W_{pq}^{E_{5b}E_{5b}}$	L5b exc	L5b exc	Local	0.12	5	18
$W_{pq}^{I_{5b}E_{5b}}$	L5b inh	L5b exc	Local	0.10	5	19
$W_{pq}^{E_{5b}I_{5b}}$	L5b exc	L5b inh	Local	0.20	3	20
$W_{pq}^{E_{5b}E_{5r}}$	L5b exc	L5r exc	Local	0.02	5	21
$W_{pq}^{E_{6r}E_{23}}$	L6r exc	L2/3 exc	Local	0.01	5	22
$W_{pq}^{E_{6s}E_{5b}}$	L6s exc	L5b exc	Local	0.08	5	23
$W_{pq}^{I_{6r}E_{6r}}$	L6r inh	L6r exc	Other	0.02	5	24
$W_{pq}^{E_{6r}I_{6r}}$	L6r exc	L6r inh	Local	0.28	5	25
$W_{pq}^{I_{FIX}E_{23}}$	Fix inh	L2/3 exc	Local	0.004	5	26
$W_{pq}^{I_{FIX}I_{5r}}$	Fix inh	L5r inh	Other	0.10	3	27

Note. exc = excitatory; inh = inhibitory; fix = fixation.

$$W_{lr} = \sum_{k=1}^{2} \gamma_k \delta_{(p-k-1)q} \qquad (8)$$

The parameter $\gamma_k = 1 - (k-1) \times 0.1$ defined the change in bias with increasing distance from the fovea. The second part was weaker and targeted populations with a retinotopic representation at the far left. This connection bias favored large return saccades.

$$W_{cr} = \sum_{k=1}^{4} \sum_{l=z}^{z+3} \beta_k \delta_{pk} \delta_{1q} \qquad (9)$$

$\beta_k = 0.7 - (k-1) \times 0.1$ decreased with increasing distance from the left end of the visual field. The sum of these two parts defined the connection from Layer 6r to Layer 4.

$$W_{pq}^{E_4E_{6r}} = w^{E_4E_{6r}}(W_{lr} + W_{cr}) \qquad (10)$$

The feedback connection from excitatory neurons of Layer 5b to excitatory neurons in Layer 2/3 [12] targeted the foveal representation only.

$$W_{pq}^{E_{23}E_{5b}} = w^{E_{23}E_{5b}} \delta_{x_{fov}p} \qquad (11)$$

This connection reset the attentional activity in Layer 2/3 back to the fovea after each saccade. The foveal inhibitory population within Layer 6r received excitatory input from all excitatory populations in Layer 6r [24] and locally inhibited the foveal, excitatory population in Layer 6r [25].

$$W_q^{I_{6r}E_{6r}} = w^{I_{6r}E_{6r}} \text{ and } W_q^{E_{6r}I_{6r}} = w^{E_{6r}I_{6r}} \qquad (12)$$

Finally, the fixation neurons received excitatory input from the foveal representation in Layer 2/3 [26] and were inhibited by the inhibitory neurons in Layer 5r [27].

(*Appendix continues*)

$$W_q^{I_{FIX}E23} = w^{I_{FIX}E23}\delta_{qx_{fov}} \text{ and } W_q^{I_{FIX}I_{5r}} = w^{I_{FIX}I_{5r}} \quad (13)$$

The fixation neurons prevented the buildup of motor activity by inhibiting all retinotopic positions in Layer 5r [17].

$$W_p^{E_{5r}I_{FIX}} = w^{E_{5r}I_{FIX}} \quad (14)$$

Implementation of the Word Processing (WP) Module

The WP module consisted of three arrays of word length detectors for the three lengths of word. We assumed such word detectors to exist and modeled them directly by changing their input according to the visual input. Note that the dorsal visual input to the FEF was feature independent (for a possible implementation of such word detectors in a network, see Dehaene et al., 2005; Riesenhuber & Poggio, 1999). A retinotopically specific release of inhibition driven by Layer 2/3 of the FEF directly selected which word detectors should fire. Words could be processed by the WP module only within the foveal and parafoveal region. Hence, the word detectors responded with a firing rate of around 70 Hz only if a word of their preferred length was at the attended location.

The population of the actual recognition neurons was driven by the word detectors. The recognition population was modeled similarly to Layer 5 in the FEF. A population of ramping neurons induced a delay between the onset of activity in any of the feature detectors and a burst of activity that signaled the recognition. Recognition in this context was equivalent to a signal that attention could be withdrawn from the currently processed word. The recognition burst excited inhibitory neurons in Layer 2/3 at the attended position and at the fovea, and therefore turned off the memory and attentional activity in Layer 2/3. Hence, the recognition signal corresponded to the command to release attention. Figure A2 illustrates how the WP module and the FEF model interacted.

As the FEF circuit, the WP module contained populations of IF neurons at 21 retinotopic positions along the horizontal axis. An array of populations of 25 inhibitory neurons (IW) at each position and three arrays of populations of 100 excitatory neurons (EW1, EW2, EW3) constituted the word length detection part of the module. External background inputs drove the inhibitory neurons to a high firing rate of around 40 Hz while the excitatory populations received less background input. The recognition part of the WP module consisted of a single population of 100 excitatory ramping neurons (ERr), one population of 100 excitatory (ERb), and 25 inhibitory (IRb) bursting neurons. An array of 21 populations of 100 excitatory projection neurons (ERp) finally sent the signal back to the FEF. The ramping to threshold in the WP module was similar to those of Layers 5r and 5b within the FEF.

The dynamics of single neurons in the WP module were the same as for the FEF (Equation 1). Neurons in ERb and ERr had excitatory neuronal parameters, and neurons in EW1, EW2, EW3, IW, IRb, and ERp had inhibitory parameters (see Table A1). The background conductances of all populations of the WP module are given in Table A4.

Table A4

Mean Values of Fluctuating Background and External Inputs

Neuronal population	μ_e	μ_i
IW	0.55	0.34
EW1, EW2, EW3	0.42	0.30
ERr	0.45	0.33
ERb	0.38	0.30
IRb	0.32	0.34
ERp	0.40	0.33
Ventral input to EW1, EW2, EW3	0.198	

The function of the individual populations was the following. The three word length detector populations EW1 (short words), EW2 (medium-length words), and EW3 (long words) received a direct word-length-dependent input that drove the neurons to a firing of around 70 Hz. These inputs were constantly turned on. However, populations in IW suppressed the firing of the EW populations. Inhibition of the firing in a particular population of IW released this inhibition and allowed the corresponding feature detector to fire at 70 Hz. The inhibition of the IW neurons was realized by a direct inhibition from Layer 2/3 of the FEF. Remember, neurons in Layer 2/3 signaled the focus of attention. Hence, the word length detectors responded only at the current focus of attention.

The recognition part of the WP module was modeled similarly to Layer 5 in the FEF, with an array of ramping and bursting neurons. All populations of word length detectors drove the ramping neurons (ERr). Bursting in the ERb population occurred when the ramping neurons crossed a threshold. Finally, a population of projection neurons ERp excited the inhibitory neurons in Layer 2/3 of the FEF. All populations of projection neurons were activated by bursts from ERb, but they could respond only at the currently attended position and at the fovea. A release of inhibition similar to that of the EW populations enabled this retinotopic projection back to the FEF. Connections within the WP module were defined as for the FEF circuit. The average weights and time constants of all connections are listed in Table A5.

Figure A2 shows the detailed architecture of the WP module and its connections to the FEF. Connections are numbered by Roman numbers and are referred to within brackets (e.g., [i] for the connection from IW to EW1). Connections from the IW populations to all EW populations were local [i, ii, and iii]. EW populations representing the fovea and the parafovea connected to the ERr population with different weights, reflecting the desired different word processing speeds for words within the fovea and parafovea [iv, v, and vi]. The connection matrix was given by

$$W_q^{ERrEW} = w^{ERrEW}\left(\sum_{x \in F} \delta_{xq} + \rho_W \sum_{x \in PF} \delta_{xq}\right). \quad (15)$$

Table A5

Synaptic Connections in the Word Processing Module and Between the Word Processing Module and Frontal Eye Field

Connection	To	From	Type	Weight	τ (ms)	Figure A2
$W_{pq}^{EW1\,IW}$	EW1	IW	Local	0.60	5	i
$W_{pq}^{EW2\,IW}$	EW2	IW	Local	0.60	5	ii
$W_{pq}^{EW3\,IW}$	EW3	IW	Local	0.60	5	iii
$W_{pq}^{ERr\,EW1}$	ERr	EW1	Other	0.00336	5	iv
$W_{pq}^{ERr\,EW2}$	ERr	EW2	Other	0.00312	5	v
$W_{pq}^{ERr\,EW3}$	ERr	EW3	Other	0.00288	5	vi
$W_{pq}^{ERr\,ERr}$	ERr	ERr	Local	0.0012	50	vii
$W_{pq}^{ERb\,ERr}$	ERb	ERr	Local	0.012	5	viii
$W_{pq}^{ERb\,ERb}$	ERb	ERb	Local	0.028	5	ix
$W_{pq}^{IRb\,ERb}$	IRb	ERb	Local	0.04	5	x
$W_{pq}^{ERb\,IRb}$	ERb	IRb	Local	0.08	3	xi
$W_{pq}^{ERr\,IRb}$	ERr	IRb	Local	0.24	5	xii
$W_{pq}^{ERp\,IW}$	ERp	IW	Other	0.60	5	xiii
$W_{pq}^{ERp\,ERb}$	ERp	ERb	Global	0.02	5	xiv
$W_{pq}^{IW\,E23}$	IW	E 2/3	Local	0.024	5	R1
$W_{pq}^{I23\,ERB}$	I 2/3	ERb	Local	0.016	10	R2

Here, $F = \{10, 11, 12\}$ are the indices of foveal populations, and $PF = \{8, 9, 13, 14\}$ are the indices of the parafoveal populations. The connections from parafoveal representations in the EW populations to the ERr population were by a factor ρ_W weaker than the ones in the fovea ($\rho_{W1} = 0.85$, $\rho_{W2} = 0.7$, $\rho_{W3} = 0.7$). Connections between the single populations of ERr, ERb, and IRb are completely defined by their weights given in Table A5.

The projection neurons ERp were globally excited by ERb. In addition, a strong local inhibition from IW allowed them to fire only when IW was suppressed.

$$W_{pq}^{ERpIW} = w^{ERpIW}(1 - \delta_{zq})\delta_{pq} \qquad (16)$$

Note that the foveal representation in ERp was not suppressed and, therefore, fired for every recognition irrespective of the current focus of attention. This foveal suppression resolved ambiguities that occurred when word recognition and the saccade occurred at the same time. In this case, the internal resetting of the attention to the fovea within the FEF was overruled by the inhibition of the WP module. Without this foveal suppression, some words would be processed twice, once in the parafovea and a second time in the fovea.

There were two connections between the FEF and the WP module. First, excitatory neurons in Layer 2/3 of the FEF inhibited the IW neurons at the same retinotopic position [R1]:

$$W_{pq}^{IWE23} = w^{IWE23}\delta_{pq}. \qquad (17)$$

Of course, it is physiologically unrealistic that excitatory neurons inhibit another population of neurons. However, instead of introducing an additional inhibitory population to "invert" the excitatory signal, we chose to allow the excitatory populations of Layer 2/3 to inhibit population IF. The response back from the WP module to Layer 2/3 of the FEF excited inhibitory neurons [R2]:

$$W_{pq}^{I23ERp} = w^{I23ERp}\delta_{pq}. \qquad (18)$$

Whenever a visual target was recognized (i.e., whenever a population in ERb burst), inhibitory Layer 2/3 neurons at the corresponding retinotopic position were excited through this connection. Table A5 lists all the parameters of the connections of the WP module.

Data Analysis

The spiking of the populations of all neurons within the network was saved for each simulation. Results are reported as population activities. The number of spikes within a population was counted in time bins of 1 ms and then smoothed by a synaptic kernel (see Sato & Schall, 2003):

$$S(t) = \frac{(1 - e^{-\frac{t}{\tau_{\text{rise}}}})e^{-\frac{t}{\tau_{\text{decay}}}}}{\int_0^\infty (1 - e^{-\frac{t'}{\tau_{\text{rise}}}})e^{-\frac{t'}{\tau_{\text{decay}}}}\, dt'}, \qquad (19)$$

where $\tau_{\text{rise}} = 1$ ms and $\tau_{\text{decay}} = 10$ ms.
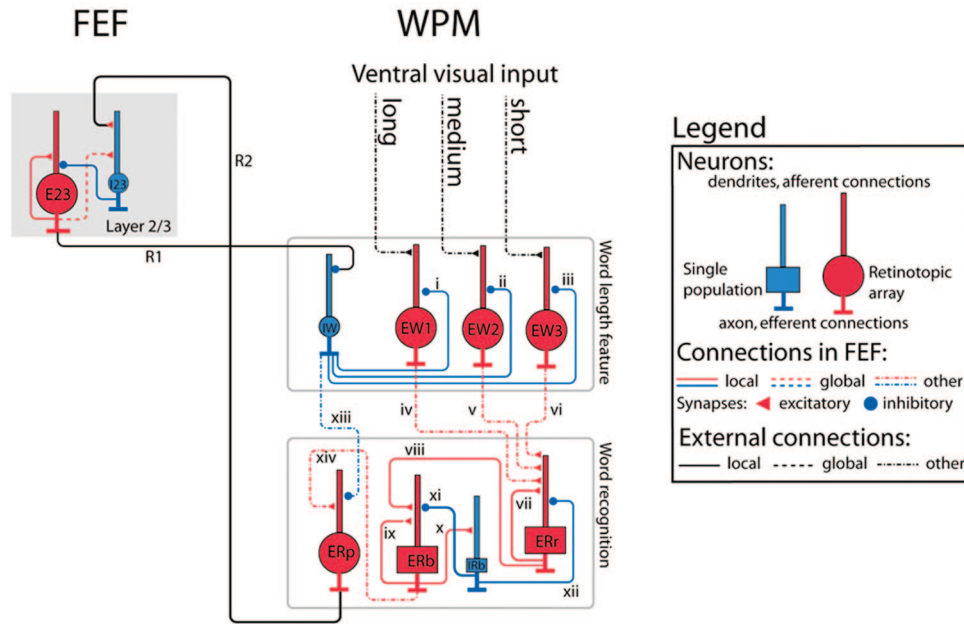
(*Appendix continues*)

Figure A2. The word processing module (WPM) and its connections to the frontal eye field (FEF). The same conventions are used as in Figure A1. Three arrays of word length feature detectors (EW1, EW2, and EW3) received ventral visual input when a word of corresponding length was in their receptive field. They responded only when the inhibition from IR was released due to the input from Layer 2/3 of the FEF (Connection R1). Different word lengths drove word recognition (ERr) at different speeds. Finally, recognition bursts in ERb excited inhibitory neurons in Layer 2/3 at the attended position via the projection neurons ERp in the WP module. A detailed list of all weights is given in Table A5.

For the behavioral analysis, saccade traces were directly given by the motor activity in Layer 5b. The allocation of attention was inferred from the activation pattern in Layer 2/3. We defined attention being allocated to a particular retinotopic position whenever the corresponding population activity in Layer 2/3 was above 30 Hz. The recognition of words was associated with the population activity of ERb in the WP module. When this activity crossed a threshold of 20 Hz from below, the attended word was considered as being recognized.