A Connectionist Multiple-Trace Memory Model for Polysyllabic Word Reading

Bernard Ans
Université Pierre Mendès France and Centre National de la Recherche Scientifique

Serge Carbonnel
Université de Savoie and Centre National de la Recherche Scientifique

Sylviane Valdois
Université Pierre Mendès France and Centre National de la Recherche Scientifique

A connectionist feedforward network implementing a mapping from orthography to phonology is described. The model develops a view of the reading system that accounts for both irregular word and pseudoword reading without relying on any system of explicit or implicit conversion rules. The model assumes, however, that reading is supported by 2 procedures that work successively: a global procedure using knowledge about entire words and an analytic procedure based on the activation of word syllabic segments. The model provides an account of the basic effects that characterize human skilled reading performance including a frequency by consistency interaction and a position-of-irregularity effect. Furthermore, early in training, the network shows a performance similar to that of less skilled readers. It also offers a plausible account of the patterns of acquired phonological and surface dyslexia when lesioned in different ways.

Much research in cognitive psychology and the neuropsychology of language is aimed toward understanding how orthography is translated to phonology when people read aloud and how this transcoding process is acquired. Among the various theoretical conceptions of the architecture of the reading system that have been described, two main classes of models have emerged, dual-route models (Coltheart, 1978, 1985; Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart & Rastle, 1994; Coslett, 1991; Funnell, 1983; Patterson & Morton, 1985; Patterson & Shewell, 1987) and parallel distributed processing (PDP) models (Hinton & Shallice, 1991; Plaut & McClelland, 1993; Plaut, McClelland, Seidenberg, & Patterson, 1996; [hereinafter PMSP96]; Plaut & Shallice, 1993; Seidenberg & McClelland, 1989 [hereinafter SM89]).

The Dual-Route Approach

The main feature of the dual-route model of reading is that it postulates the existence of two different processing routes for converting print to sound. These are a lexical route allowing access to stored knowledge about familiar (previously learned) words and a separate nonlexical route involving print-to-sound mapping rules. It is hypothesized within this framework that the phonological and orthographic representations of all previously learned words are stored in a mental lexicon in long-term memory. By this view, reading aloud by the lexical pathway is achieved by first accessing the stored orthographic representation corresponding to the input word and then retrieving the whole phonological representation of the word. According to some hypotheses, the lexical pathway is further subdivided into a semantic route and a nonsemantic route. The lexical semantic route allows access to semantic information from orthographic representations, prior to phonological retrieval, whereas the nonsemantic lexical pathway allows the phonological representation to be directly addressed from orthography without semantic mediation. Only previously learned words (that have entries in the mental lexicon) can be successfully read aloud by the lexical procedure. Reading aloud of items that have no entries in the mental lexicon (i.e., new words or pseudowords) can only be performed via the nonlexical pathway. This route depends on the use of a set of general rules specifying correspondences between orthography and phonology. The nonlexical procedure typically derives pronunciations that obey standard orthography-to-phonology correspondences. It will thus produce incorrect pronunciations (i.e., regularization errors) when the input is an irregular word. Thus, the correct pronunciation of irregular words can be obtained only through the lexical pathway by accessing the appropriate stored representation in the orthographic and phonological lexicons, whereas the correct pronunciation of regular words can be achieved by either the lexical or nonlexical routes. The dual-route model further postulates that processing orthographic input starts simultaneously in the
two routes and works in parallel whatever the nature of the input letter string (regular word, irregular word, or pseudoword) because there is no way of determining, prior to lexical access, whether an orthographic stimulus is a familiar word or a new word. As a consequence, for all words that an individual has already learned, an "addressed" phonological representation would be delivered by the lexical route and an "assembled" pronunciation would be computed by the nonlexical route. In their dual-route cascaded (DRC) model (a computational version of the dual-route model), Coltheart et al. (1993) claimed that phonological information from both the lexical and nonlexical pathways converges toward a common component of the model, the phoneme system, which comprises a set of units representing all possible phonemes and their position. These units are therefore activated by inputs from both pathways. The DRC model also incorporates assumptions about the time course of the two processing routes. Lexical processing is viewed as being typically faster than nonlexical processing. However, nonlexical processing is not so slow that its input never reaches the phoneme level when the two routes simultaneously produce different outputs (alternative pronunciations in a given position) as is the case for low-frequency irregular words.

A wide range of experimental findings from skilled readers (Beauvois & Derouesné, 1979; Coltheart & Rastle, 1994; Monsell, Patterson, Graham, Hughes, & Milroy, 1992; Paap & Noel, 1991) as well as neuropsychological evidence from studies on acquired reading disorders (Coltheart, Sartori, & Job, 1987; Marshall & Newcombe, 1973; Patterson, Marshall, & Coltheart, 1985) have been interpreted as providing support for the existence of two distinct procedures for translating print to sound (for discussion, see Carbonnel & Ans, 1996; Humphreys & Evett, 1985). The dual-route theory in particular accounts for the regularity by frequency interaction that has been observed in a number of naming latency studies (Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Trabaran & McClelland, 1987). Experimental studies indeed showed longer naming latencies for irregular words than for regular words if they are of low frequency but no such effect for words of high frequency. Within the DRC model (Coltheart et al., 1993), this regularity effect is interpreted as resulting from competition between the lexical and nonlexical phonological outputs in the following way: The rate at which activation rises throughout the lexical route is a function of word frequency so that the more frequent a word is, the sooner its corresponding phonemes will be activated in the phoneme system. As a consequence, activation of this system by the lexical route will be completed for high-frequency words (be they regular or irregular) before substantial activation (if any) gets to the system from the nonlexical route. In this case, irregularity does not affect performance. In contrast, for low-frequency words, lexical and nonlexical outputs simultaneously activate units in the phoneme system. This will result in no latency cost for low-frequency regular words because phonemes activated by the two sources will be the same but will generate conflicts whenever the stimulus is a low-frequency irregular word. In this case, the time taken to resolve the conflict will affect naming latency so that only low-frequency irregular words will suffer a latency cost (see Coltheart & Rastle, 1994).

The dual-route model further provides a straightforward explanation of two main forms of acquired dyslexias (reading disorders caused by brain damage in a previously literate individual). The pattern of surface dyslexia (Coltheart, Masterson, Byng, Prior, & Riddoch, 1983; Safran & Marin, 1977; Shallice, Warrington, & McCarthy, 1983) that is characterized by a selective impairment of irregular word reading, resulting in the production of regularization errors, is viewed as reflecting damage to the lexical route and preservation of the nonlexical pathway (Coltheart et al., 1993). In contrast, the pattern of phonological dyslexia (Beauvois & Derouesné, 1979; Derouesné & Beauvois, 1985; Shallice & McCarthy, 1985; Shallice & Warrington, 1975) that is characterized by a selective pseudoword reading impairment is interpreted as resulting from damage to the nonlexical route and relative preservation of the lexical pathway (Coltheart et al., 1993).

The theoretical conception of the reading system developed within the framework of dual-route models has been challenged by models that assume that pronouncing irregular words and pseudowords does not require separate lexical and sublexical procedures. Analogy models of the reading system (Glushko, 1979; Kay & Marcel, 1981; Marcel, 1980) represented the first radical departure from the dual-route approach, postulating that the pronunciation of regular words, irregular words, and pseudowords is assigned by analogy with and by specific reference to known lexical items, novel words being pronounced by generalization from existing words. The analogy theory assumes that only whole word phonology is stored in long-term memory and that orthography-to-phonology conversion rules only exist implicitly in the integrated activation of words (Glushko, 1979). However, the ability of an implemented network based on the analogy theory to actually reproduce the basic effects that characterize skilled word and nonword reading has not been demonstrated. Another conception of the reading system has been developed within the framework of multiple-level models (Shallice & McCarthy, 1985; Shallice et al., 1983). These models also postulate, apart from the semantic pathway, a single nonsemantic route, the phonological pathway, fully competent for reading all kinds of letter strings. The phonological route is described in this class of model as incorporating spelling-sound correspondences ranging from single graphemes and phonemes to word bodies and entire morphemes so that irregular word pronunciation could be achieved on the basis of morphemic correspondences, whereas pseudoword reading would typically involve correspondences between smaller graphemic units (see Norris, 1994, for an implementation of a multiple-level model).

Another alternative to the dual-route model, assuming (apart from the semantic pathway) a single-route mechanism for read-

---

1 Coltheart et al. (1993) referred to the novel-word pronunciation system implemented by Sullivan and Damper (1992) as an explicit model of reading by analogy. Sullivan and Damper themselves claimed that their model gives an example of how the analogy processes could be implemented. However, their model differs somewhat from the analogy theory. For example, a system of grapheme-phoneme correspondence (GPC) rules is implemented that provides "a first plausible pronunciation of the letter string" to be further refined at the analogy stage" (p. 183). The precompilation of a set of GPC rules seems in disagreement with the very basis of analogy models.
ing both words and pseudowords was proposed by SM89, who developed a PDP connectionist network that learns to map the orthographic representations of words onto corresponding phonological representations.

The PDP Connectionist Approach

The SM89 general lexical processing model is composed of three layers of units, an orthographic, a phonological, and a semantic layer, and three sets of hidden units mediating between them. The units are fully connected from one layer to the other. Information is represented in terms of distributed patterns of activity over these separate groups of units, and processing is mediated by connections among the units. Knowledge of the relations among the orthographic, phonologic, and semantic patterns is encoded by the weights on connections between units. The model postulates a phonological pathway by which the appropriate phonological pattern is computed from the orthographic pattern activated by the input and a semantic pathway allowing a word to be pronounced by means of a computation from orthography to meaning and then from meaning to phonology. In fact, only a part of this general model was implemented in the connectionist network by SM89. The implemented model only simulated interactions between the orthographic and phonological layer and contained no semantic representations. Performance was determined by the connection weights, which were progressively adjusted during the training phase, using the back propagation learning algorithm of Rumelhart, Hinton, and Williams (1986). After training on a database of 2,897 monosyllabic English words, the network was able to generate the correct pronunciation of 97.3% of previously learned real words. It also succeeded in simulating a number of the basic effects characteristic of human skilled word reading, but its performance on pseudoword reading remained far below that of skilled readers (Besner, T威尔, McCann, & Seergobin, 1990; Plaut et al., 1996). When lesioned, the network also failed to show a pattern of performance similar to surface dyslexia (Patterson, Seidenberg, & McClelland, 1989). To overcome these limitations, PMSP96 (see also Plaut & McClelland, 1993) developed new connectionist networks that kept the assumptions of the original SM89 model but adopted new representational schemes that were better able to capture the regularities between orthography and phonology. When trained on a database similar to that of the SM89 model (2,998 monosyllabic words), the networks were able to accurately name both monosyllabic (regular and irregular) words and pseudowords at a level of performance very similar to that of skilled readers. PMSP96 (see also Plaut & McClelland, 1993) therefore clearly demonstrated the capacity of a system exposed only to the orthographic and phonological forms of words to generate the correct pronunciation of both irregular words and pseudowords. The PMSP96 simulations are clear evidence against the claim that the correct pronunciation of pseudowords can be achieved only by the use of an explicit set of pronunciation rules. The networks further exhibited the expected regularity by frequency interaction observed in skilled readers. Following SM89, the notion of regularity was replaced by a more general notion of consistency. A word is said to be consistent if its pronunciation agrees with those of similarly spelled words (its orthographic neighbors). Regularity effects are then viewed as a particular class of consistency effect that arises in the network from the influence of all learned items. However, the network (described in Simulation 3) restricted to the implementation of a fully competent phonological pathway again failed, when lesioned, at reproducing the pattern of surface dyslexia (in particular that of the most severe forms of surface dyslexia). This failure was interpreted as resulting from a limitation of the implemented network that did not take into account a semantic contribution in reading. PMSP96 argued that surface dyslexia should better be viewed as reflecting the behavior of an undamaged but isolated phonological pathway that had learned to map orthography to phonology in the context of support from semantics. Then, a new network (Simulation 4) was trained that incorporated a graded division of labor between the semantic and phonological pathways.

The fourth PMSP96 implementation took the form of a feedforward network consisting of three layers of units: an input layer of 105 grapheme units, an output layer of 61 phoneme units, and an intermediate level of 100 hidden units. The contribution of the semantic pathway was only approximated by adding some amount of external input from the putative semantic pathway to the phoneme units. The phonological pathway was trained with a concurrently increasing contribution from semantics. It was first demonstrated that the network trained with semantics still exhibited a performance similar to that of skilled readers. Furthermore, the network exhibited the pattern of surface dyslexia when the semantic input was disrupted. This pattern therefore resulted from the normal operation of the intact phonological pathway when deprived of additional semantic activation. The pattern of phonological dyslexia was not simulated because it was viewed as resulting from either selective damage of the phonological pathway—so that reading would be primarily accomplished by a semantic pathway fully competent in familiar word reading (this pathway was not implemented in the PMSP96 work)—or phonological damage per se.

A New Theory of the Reading System

The dual-route model and the PDP connectionist approach differ on two main aspects. First, the central dogma of the dual-route model is that pseudoword reading requires a system of conversion rules (Coltheart et al., 1993; Reggia, Marsland, & Berndt, 1988), whereas the key feature of the SM89 and PMSP96 model is that all types of letter strings, including pseudowords, can be read solely on the basis of word knowledge without stipulation of a separate rule system. A second difference that was typically viewed as being intimately related to the first one concerns the reading procedures. The PDP approach postulates that the ability to accurately pronounce all kinds of letter strings is supported by a single uniform procedure, whereas the dual-route model claims that two separate (lexical and sublexical) procedures are required.

With respect to the former point, the claim of PDP models seems to go beyond that of dual-route models in not postulating the need for any metaknowledge about the way orthography relates to phonology. The idea that the ability to accurately read any kind of input string is determined only by the properties of all the words (in particular its neighbors) seems more parsimonious and appears to provide a more direct account of the consis-
tency effects among words. With respect to the second point, however, a number of empirical results have been interpreted as providing support for the existence of two distinct reading procedures (Monsell et al., 1992; Paap & Noel, 1991; Seidenberg et al., 1984; Weekes, 1997). Carbonnel, Valdois, Bozon, Martinet, & Ans (1998) more directly addressed this question by contrasting data on French polysyllabic word and pseudoword reading. The authors investigated the effects of the length of printed words and pseudowords (in terms of number of syllables) on naming latency and on the presentation time minimally required for the accurate reading of these letter strings. In naming, similar latencies were obtained for mono- and bisyllabic words with only a marginally significant increase in naming latency for tri-syllabic words. In contrast, a highly significant length effect characterized pseudoword naming, and the length by lexical status interaction was highly significant. Essentially similar results were obtained in a threshold measure for words than for pseudowords. More crucially, increased for pseudowords as providing support for the existence of two distinct reading procedures is particularly clear for polysyllabic items. A new model of the reading system that accounts for a broad range of phenomena related to the processing of not only monosyllabic but also polysyllabic words and pseudowords is described here. The model goes beyond the earlier models of poly-syllabic word and pseudoword reading. The patterns of performance characteristic of these two forms are processed as a whole. The items that are not recognized as familiar cannot be processed globally. Processing will then apply to the largest initial component of the printed letter string that can be recognized as familiar by the system and will progress by shifting from one familiar spelling pattern to another up to the end of the string.

Finally, in contrast to PMSP96's view, the current model assumes the existence of a phonological pathway fully competent for reading all kinds of letter strings, thus minimizing the role of semantics in normal single word reading and consequently in the explanation of surface and phonological dyslexia. The patterns of performance characteristic of these two forms of dyslexia are replicated here after specific damage within the phonological pathway itself.

The model is in keeping with the "episodic" theories that have been particularly influential in the field of cognitive psychology in the course of the last decade. These theories assume that no general knowledge is abstracted from learning episodes. Learning only consists of adding new separate episodic traces in memory without modifying memory structure. Knowledge evoked by a given current episode arises at the time of retrieval only through the blending of multiple activated episodic traces. This approach has been adopted in the domains of memory (e.g., Hintzman, 1988; Jacoby & Brooks, 1984), object meaning (Carbonnel, Charnillet, David, & Pellat, 1997), categorization (e.g., Hintzman, 1986; Kruschke, 1992; Medin & Schaffer, 1978; Nosofsky, 1984; Nosofsky & Palmeri, 1997), judgment (Kahneman & Miller, 1986), automatization (Logan, 1988; Nosofsky & Palmeri, 1997), problem solving (Ross, 1984), and social judgments (Smith & Zarate, 1992). It is shown here that the multitrace episodic approach can also be successfully applied to reading.

The Model

The connectionist reading network (Figure 1) is composed of four layers of neuronlike processing units: an orthographic
input layer O1, a second orthographic layer O2 having the same structure as O1, a phonological output layer P, and an intermediate layer EM (the "episodic memory layer") mediating between them. Each O1 unit is connected to each EM unit, and each EM unit in turn is connected to each O2 unit and each P unit. Also, two other input patterns send activation to all EM units: the "reading mode" pattern, denoted RM, associated with the current processing mode (global or analytic) and a pattern representing the state of the context surrounding the current orthographic input.

**Orthographic Input Layer**

As shown in Figure 2, the orthographic input layer O1 is composed of a set of unconnected clusters of elementary units, each cluster corresponding to a specific position in an input letter string (a maximum number of 21 clusters was adopted in the simulations). Each cluster is made up of 41 units specifically coding for the 40 French alphabetic characters plus one virtual character (denoted #) coding for word boundaries, giving a total of 861 O1 units. The coding of orthographic input over O1 depends on the status, entirely or partially focal, that is attributed to this input as a function of its processing mode (as described below). The coding of the input is also dependent on the determination of a focalization point (FP) within a focal window.

When the input is processed in the global mode (see Figure 2a), all its orthographic characters, including left and right boundaries, belong to a so-called focal window. In contrast, in the analytic mode (Figure 2b), the focal window is restricted to a part of the input letter string, the other letters being coded as contextual. The maximum number of left and right contextual characters was limited in the simulations (this number was determined by the context size parameter cs).

Letter position encoding is performed by reference to an origin cluster (labeled O cluster) aligned in front of the FP. The location of this FP is determined by the initial detection of an orthographic vocalic "motif" within the current focal window. The orthographic vocalic motifs (e.g., a, ai, an, ain, and eau) are defined as functional sequences, or vocalic graphemes, of one, two, or three letters. These graphemes typically correspond in French to a single vocalic phoneme (e.g., /a/, /e/, /β/, /ε/, and /o/, respectively) except for some that correspond to a sequence of a "glide + vowel" (e.g., oi ← /wa/). In both the global and the analytic processing modes, the FP is located in front of the rightmost letter of the first vocalic grapheme (underlined) of the focal window. The alignment of the 0 cluster with the FP determines the O1 layer activity: The units that respond to focal characters (black units) have an activity equal to 1, whereas those responding to contextual characters (shaded units) have an activity equal to ca, with ca < 1. (The empty units denoting nonactive units). (a) The French word *fourmi* (ant) processed in the global mode in which the focal window includes the whole word plus the boundary characters. (b) The French word *caméra* processed in the analytic mode in which the focal window is restricted to one of its syllables; the number of contextual letters is taken as equal to three (cs = 3) for purposes of illustration.

---

3 From now on, the term orthographic character is used to refer to both a letter or the virtual character #.

4 In the global mode, the FP typically points to characters belonging to the beginning of the input word; this is thus compatible with the notion of convenient viewing position (Ellis, Flude, & Young, 1987; O'Regan, 1981; O'Regan, Lévy-Schoen, Pynte, & Brugallère, 1984; Rayner & Pollatsek, 1987).
Perceptual Identification section for a justification of this assumption. For example (see Figure 2a), when the French word *fourmi* (*ant*) is processed in the global mode, the FP will be located towards the right-most letter *u* of the first vocalic grapheme *ou* detected from the beginning of the focal window (which covers in this case the whole word).

The implementation of the early perceptual stages leading to the recognition of vocalic graphemes is beyond the scope of the current work. It is assumed that specific visual analysis procedures for the detection of vocalic graphemes are developed during the course of reading acquisition. Different stages of perceptual learning can be distinguished from the detection of visual features to the construction of letter codes and the recognition of graphemic units corresponding to particular letter combinations (cf. Laberge & Samuels, 1974). The present model simulates the reading procedures that take place after early perceptual analysis at a stage at which graphemic units can be extracted.

Following the positional encoding procedure described above, each input character is mapped onto a positional cluster from the 0 cluster position. Each character of the orthographic input is coded by a single unit selectively activated in the corresponding O1 cluster: Units responding to focal characters (the characters belonging to the focal window) have an activity equal to one, and activity of the units coding for contextual characters is reduced to a value noted *ca* (*ca* < 1), which is a model parameter.

**Phonological Layer**

Phonological encoding on layer P takes into account three kinds of natural phonological units: phonemes, syllables, and syllabic constituents (onset and rime). The role of the phoneme as a basic representational and functional unit is widely recognized in models of linguistic competence (Chomsky & Halle, 1968) as well as in psychological accounts of normal or pathological performance (Blumstein, 1990; Dell, 1986, 1988; Shattuck-Hufnagel, 1979; Valdois, Joannette, & Nespoulous, 1989; Valdois & Nespoulous, 1994, for a review in the field of neuropsychology). A common assumption is that spoken words are strings of phonemes organized into syllables. The concept of the syllable was first developed in linguistic theory (Halle & Vergnaud, 1980; Kaye & Lowenstamm, 1984; Selkirk, 1982), where the syllable was described as having a hierarchical internal structure. It consists of two major constituents, the onset and the rime, the latter being further subdivided into a nucleus and a coda. In a monosyllabic word, the onset is the consonantal phoneme (e.g., /m/ in /ềm/; *sea*) or the consonantal cluster (/pr/ in /pre/, *meadow*) preceding the vocalic phoneme; the rime is the vowel and any following consonants (e.g., /ask/ from /park/; *garden*). The vocalic phoneme of the rime is called the nucleus, the remaining consonantal segments constituting the coda. The relevance of the syllable as a functional unit has also been demonstrated in numerous studies concerned with spoken word processing (Blumstein, 1978; Segui, 1984), and evidence for onset—rime coding has been provided by many psychological and neuropsychological findings (Treiman & Chaftet, 1987; Treiman & Danis, 1988; Treiman, Goswami, & Bruck, 1990; Valdois, 1990).

The phonological layer P (Figure 3) consists of a set of clusters, each made up of 39 units specifically coding for the 38 French phonemes plus one silent phoneme (*/*). Layer P is divided into a number of cluster subsets potentially coding for the successive syllables of a phonological word. Within each syllabic subset, each cluster corresponds to a phonemic position determined by reference to the syllable nucleus (the 0 cluster). Negative positions code the onset of the syllable, the zero and positive positions corresponding to the rime. The phonological encoding of the bisyllabic French word *fourmi* is shown as an illustration. Syll = syllabic subset.

**Learning the Pronunciation of a Printed Word**

Figure 4 depicts how orthography-to-phonology correspondences are acquired by the network during learning. An activity

---

4 Representational units smaller than phonemes were not considered.
pattern is generated over the O1 layer from the input letter string according to the previously defined coding procedure. At the same time, the phonological counterpart is made available, which gives rise to an activity pattern over the phonological layer P. Learning can be done in the global or analytic mode. When a word is learned in global mode (Figure 4a), all its orthographic characters (including left and right boundaries) belong to the focal window and its phonological form is simultaneously provided over the phonological layer. When it is learned analytically (Figure 4b), each syllable of the word is succes-
sively presented as input over O1 (with the orthographic context of the word it originates from) while its phonological counterpart is made available over the first syllabic subset in P. Learning a given word (e.g., \textit{caméra}) within this latter mode will proceed from left to right, each syllabic correspondence being successively learned: \textit{#camér} → /ka/\textit{, #caméra#} → /ne/\textit{, améra#} → /ra/, with a context size parameter \( c \) equal to 3 in this example.

In the general model (see Figure 1), the layer EM is in part activated by the pattern generated over the O1 input layer and by another input activity pattern \( RM \), which is associated with the processing mode of the current orthographic object. The size of the \( RM \) vector may be chosen arbitrarily; however, a two-component vector is sufficient here because \( RM \) can take only two orthogonal states: G when the orthographic input is processed in global mode and A when it is processed in analytic mode (\( G = \{ 10 \} \) and \( A = \{ 01 \} \)). An “orthographic event” is represented by the current O1 pattern and its associated processing mode \( RM \).

In learning, the EM layer can keep separate and specific memory traces of occurring episodes (Hintzman, 1984, 1986, 1988). An episode is defined here as being formed of an orthographic event plus the current state of an environmental context (the sentence, and more generally the situational context, surrounding the input orthographic object). This environmental context, which is likely to fluctuate continually, can take a priori an unlimited number of unknown states. However, it will be assumed that the environmental context is a random entity and that the network has a limited sensitivity to contextual fluctuations so that only a limited number of contextual states can be distinguished by the network.

The model postulates that creating separate EM traces during learning is governed by the following basic principle: Only new episodes (never yet processed) give rise to additional separate EM trace, so the episodic memory layer EM keeps separate traces of the distinguishable episodes alone (this is slightly different from the Hintzman’s multitrace principle stipulating that each experienced episode creates a new separate trace). Thus, any new orthographic event would always induce a new EM trace, whatever its environmental context (the corresponding episode being then necessarily new). In contrast, the presentation to the network of a previously processed orthographic event would lead to the creation of an additional separate EM trace only when the current state of the environmental context is distinguishable from any of the contexts that have been previously associated with this same event (hence forming a new episode). In other words, a separate trace will be created each time the same orthographic event is encountered in a distinguishable context. The number of traces created for a repeated orthographic event is specified in the \textit{Word Frequency Account} section.

In connectionist terms, localized EM traces would be created as follows. Most of the O1-EM and RM–EM connections initially have random weak weights that have not yet undergone any strengthening. When the system is learning, a single EM unit can be activated by the current episode. If this episode is a new one, then this unit is selected among those with not-yet-strengthened connection weights. The O1 orthographic input pattern and the RM pattern are then both stored in the weights of the connections originating from O1 and RM toward the only active EM unit. Then, the set of the newly strengthened weights (the “synaptic” weight vector) supported by this selected EM unit constitutes the specific episodic trace of the orthographic event. When an already processed episode (same orthographic event and same context) is encountered again, the EM unit specifically coding for this episode is selected again and its memory trace remains unchanged. The EM traces that are created by orthographic events related to whole words (global mode) are called \textit{word traces}; those induced by events related to focal segments in their orthographic context (analytic mode) are termed \textit{segment traces}.

The setting of the connection weights \( \mu_q \) from O1 units \( j \), with an activity \( e_j (e_j = 0, 1, \text{ or } c) \), to the selectively activated EM unit \( i \) (activity equals one) is governed by the following learning rule:

\[
\mu_q = e_j / \sum e_i,
\]

where the summation index \( i \) ranges over all the O1 units. Notice that the sum of the input connection weights \( \mu_q \) supported by a given episodic unit \( i \) equals one. In the same way, the RM pattern, whose components are denoted \( r_j \) (with \( r_j = 0 \) or \( 1 \)), is stored in the connection weights \( \rho_q \) that link the two \( r_j \) components to the same active EM unit \( i \):

\[
\rho_q = r_j / \sum r_i,
\]

where \( i \) ranges over the two RM components. In fact, \( \rho_q \) always reduces to 0 or 1.

The EM layer transmits its activity to all the units of the O2 and P layers. The layer O2 has the same structure as O1 and receives one-to-one connections from this layer (see Figure 4). These specific connections play the role of filters in such a way that the only available information over O2 is that coming from O1’s focal window. This focal pattern over O2 consists only of teaching signals that do not contribute to the effective activation of the O2 units but serve to select the connections to be strengthened from the EM layer to the O2 layer. Within the O2 layer, the learning rule consists in setting at a value 1 the weights \( \omega_{ji} \) of the connections originating from the single active EM unit \( i \), coding for the current orthographic event, to the O2 units \( j \) that are selected by the teaching signals giving the focal part of the orthographic input. For the O2 units \( j \) that are not selected, \( \omega_{ji} = 0 \).

A similar learning rule applies at the same time within the phonological layer P. Connections from the single active episodic EM unit \( i \) toward the only \( j \) units of P that are active get

---

5 In fact, the finer scale dynamical processes that would detect new episodes and give rise to separate episodic EM traces have not been implemented here. In his ART model, Grossberg (1987a, 1987b) has proposed a neural implementation of localized memory based in particular on a mechanism for the detection of old and new input events. This detection was under the control of a vigilance mechanism that modulated the resolving power of the model, that is, the number of distinguishable input events.

6 A simple way of modeling the filter system might consist in introducing interunits with a threshold \( h (c < h < 1) \) between layers O1 and O2.
weights $\pi_{ji} = 1$, whereas all other connections between unit $i$ and $P$ have weights equal to 0.

Thus, orthography-to-phonology correspondences (for both words and segments) are learned and consequently stored in the connection weights linking the four layers of the network.

**Naming a Letter String**

During the testing phase, the network generates an activity pattern over the phonological layer in response to an orthographic input. This pattern will be taken as the pronunciation of the current input letter string.

**Layer EM Activity**

A given input, activating the episodic memory layer EM, is composed of three parts: an input orthographic string (coded over the O1 layer), a reading mode pattern $RM$ (whose state is determined according to the current reading procedure, which is specified later), and the present state of the environmental context. Contrary to the learning phase, in which layer EM activity was highly contrasted (only one of its units was active), in the testing phase this activity is only partially contrasted (several units show an output activity). The working principle of the EM layer is inspired by Hintzman’s (1984, 1986, 1988) fundamental hypothesis postulating local processing during learning and spreading-parallel processing at retrieval.

The activity of EM units varies as a function of the match between their specific memory trace and the current three-component input. First, the input activation from the O1 layer, denoted $u_i$, to a given EM unit $i$ is the sum of the activities $e_j$ of all the O1 units, each being weighted by $\mu_{ij}$ as expressed in Equation 1, which is written $u_i = \sum_j \mu_{ij} e_j$. In fact, this input activation (with $0 \leq u_i \leq 1$) to EM unit $i$ simply reflects the orthographic similarity between the orthographic trace specific to this unit and the input letter string.

Second, the input activation from the environmental context would be in principle added to the orthographic input activation. This activation would then be a function of the match between the current context and the stored one. Nevertheless, this contextual component was not implemented in the current simulation. Indeed, the network’s performance was compared only with that of human participants reading isolated words and pseudowords. Consequently, the test context could be viewed as being somewhat atypical. Hence, this context would be expected to only weakly (and not differentially) match the ecological contextual components of the EM traces that would have been stored during learning. It was therefore reasoned that the artificial test context would have no significant effect on EM activity. For this reason, the contextual component of each episode was not required to be explicitly stored in the EM traces during the learning phase.

Third, the input activation from the $RM$ pattern, denoted $v_i$, to EM unit $i$ is the weighted sum (weights $p_{ij}$ given in Equation 2) of the $RM$ components $r_j$, which gives $v_i = \sum_j p_{ij} r_j$.

Hence, only the two EM input activations originating from the input orthographic event ($u_i$ from O1 and $v_i$ from $RM$) will be taken into account in the reading tests performed in the current network. These two activation components act multiplicatively so that the resulting activation, denoted $a_i$, of an EM unit $i$ is written as

$$a_i = u_i v_i.$$  \hspace{1cm} (3)

In Equation 3, factor $v_i$ can take only two values, one or zero: $v_i = 1$ when the input $RM$ corresponds to the $RM$ that has been associated with the EM unit $i$ during learning, and $v_i = 0$ otherwise. In sum, EM word traces alone or EM segment traces alone are activated (and hence recruited in the EM layer) depending on whether the input $RM$ is in state $G$ or $A$.

During reading tests, the output activity, denoted $s_i$, of an EM unit $i$ is governed by a partial contrast-enhancement process that was simulated using the simple rule

$$s_i = a_i / \sum_i a_i,$$  \hspace{1cm} (4)

where $\gamma$ (with $\gamma > 1$) is a parameter whose value modulates the contrast enhancement of the activations. The summation index $k$ ranges over all the EM units, thus resulting in normalization of the total activity of layer EM ($\sum_k s_k = 1$). Equation 4 is similar to the one proposed by Hintzman (1984, 1986, 1988) except for the normalization property that is added here (cf. Grossberg 1973, 1988, for a detailed neural network implementation of partial contrast with normalization; see also Carpenter & Ross, 1995, where the same power rule as Equation 4 is used for contrast enhancement).

**Layer O2 and P Output**

In response to an input orthographic event, the EM layer generates in parallel an orthographic echo to the O2 layer and a phonological echo over the P layer in the following way. The O2 and P clusters are winner-take-all (WTA) clusters (Ans, 1990a, 1990b; Ans, Coiton, Gilbodes, & Velay, 1994; Grossberg, 1973, 1987a, 1987b, 1988; Kohonen, 1988, 1993, 1994; McClelland & Rumelhart, 1981), in which a single unit alone can be active at the end of a competitive process (within a WTA cluster, each unit is self-activated and inhibits all other units).

The input activation $o_j$ of each O2 unit $j$ is given by $o_j = \sum_i \omega_{ij} s_i$, where $\omega_{ij}$ are the connection weights (with value 0 or 1) from EM units $i$ to the O2 unit $j$ that have been strengthened during learning (with $o_j \leq 1$ because the total activity of layer EM equals one). In the same way, the input activation $p_j$ of each P unit $j$ is given by $p_j = \sum_i \pi_{ij} s_i$ (with $p_j \leq 1$), where $\pi_{ij}$ are the connection weights (with the values 0 or 1) from EM units to the P unit $j$. In other words, activation of the O2 and P layers arises from the parallel summation of all previously learned orthography-to-phonology correspondences, each being weighted by the activity level of its associated specific EM unit.

The O2 units (and P units) have a threshold $\theta_0$ (and, $\theta_\gamma$, respectively) so that they can get an (nonzero) initial output activity only if their input activation is equal to or above the threshold, that is, if $o_j \geq \theta_0$ (or if $p_j \geq \theta_\gamma$, respectively). According to their initial output activity, the O2 (or P) units eventually compete for the same position within a WTA positional cluster and reach a final stabilized state. The dynamical process leading to this final state, in which a single unit alone can be active (the winning unit), was not implemented in the
Two Reading Procedures

During the reading test of an orthographic object (a previously learned word, a new word, or a pseudoword), the network was always initialized in the so-called global reading procedure. In this procedure, the orthographic input is held as a whole by the O1 layer while its associated reading mode RM is initialized in the state RM = G. As during learning, in the global reading mode the whole orthographic object (including the left and right boundaries #) forms the focal window, which is positioned so that the O1 positional 0 cluster codes for the rightmost letter of the first vocalic grapheme.

In response to this input stimulus, orthographic and phonological echoes are generated over the O2 and P layers on the basis of the EM word traces alone as previously described. A matching check is then performed, in which the O2 orthographic echo is compared to the O1 orthographic input. If the O2 orthographic echo is strictly identical to the O1 input—that is, the same characters occur in the same positions—then the phonological echo is taken as the pronunciation (computed in only one cycle) of the whole orthographic stimulus. The input stimulus is then said to have been read in global mode (in a feedforward way).

When the orthographic echo differs from the O1 input (i.e., when the global matching check fails), the phonological echo is inhibited with the result that the input cannot be processed in global mode. The network then shifts to the so-called analytic reading procedure. In this procedure, the reading input mode is always RM = A so that only EM segment traces are recruited. The orthographic object is then sequentially processed, segment by segment (typically syllable by syllable), each segment being extracted by the system itself at each processing step (as described below).

A single processing step in the analytic reading mode can be summarized as follows. At the beginning of the step, the orthographic input is typically composed of two parts. The left part already having been read during preceding steps is now considered as contextual (within the limits of the context size parameter cs); the remaining right part that remains to be read is processed as focal. The focal part is here again positioned in such a way that the O1 positional 0 cluster codes for the rightmost letter of its first vocalic grapheme. In response to this current input, the orthographic and phonological echoes are simultaneously generated over the O2 and P layers; the orthographic echo is then compared, as in the global reading procedure, with the O1 orthographic input. However, in the analytic reading procedure, the O2 echo no longer must be identical to the O1 input; rather, the O2 echo has only to be strictly identical to its O1 homologue part (the same characters in the same positions). If this partial matching succeeds, the phonological echo that has been generated over P is taken as the phonological output of this current step. At the same time, the O2 orthographic echo is used to tag the corresponding identical segment of the current O1 focal window that has just been processed. The tagging is also used for specifying the remaining part of the orthographic input object to be read at the next step. This procedure is sequentially reiterated until the input orthographic object is entirely translated. At the initial step of the iterative procedure, the input contextual part is limited to the left boundary (#), the remaining part of the orthographic input being focal (see Figure 5 for an illustration of this procedure).

If, at a given step of the analytic reading procedure, the orthographic matching check does not succeed, then the corresponding phonological output is inhibited and the system temporarily enters, only for this failing step, into a finer grained orthographic analysis. In this case, the reading mode RM remains unchanged.

1 However, the dynamical process subserving WTA clusters is taken into account later in the naming latency simulations.

Figure 5. Sequential analytic reading (in three consecutive steps) of the pseudoword touspiral showing in particular that orthographic syllabic segmentation is automatically performed by the reading network itself. At the beginning of each step, the O1 orthographic input is typically composed of two parts: The right part that remains to be read is processed as focal (the O1 focal window is marked with bold letters and the focalization point FP is represented by a bottom-up arrow); the left part already read during the preceding steps is processed as contextual (italic letters, within the limits of the context size parameter cs, taken as equal to 2 for the purpose of illustration). After phonological and orthographic echoes (P and O2) are generated, and subject to the success of the partial matching test between the current O1 focal window and the O2 echo (as is always the case for the three steps in this example), the O2 echo tags (top-down arrow) the corresponding identical segment within the current O1 focal window (framed letter string). This O1 tagged segment thus corresponds to the part of the input (typically a syllable) that has been "recognized" and read (P echo) by the network. The tagging process also defines the left boundary of the remaining right part of the orthographic object to be read at the next step.

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>P Echo</td>
<td>/tu/</td>
<td>/pi/</td>
</tr>
<tr>
<td>O2 Echo</td>
<td>touspiral#</td>
<td>ompiral#</td>
</tr>
<tr>
<td>O1 Input</td>
<td>#toupiral#</td>
<td>oumpiral#</td>
</tr>
<tr>
<td>FP</td>
<td>FP</td>
<td>FP</td>
</tr>
</tbody>
</table>
(RM = A), but the current focal window is restricted to the first (leftmost) grapheme of the part of the input letter string that remains to be read. The orthographic context is not taken into account (i.e., parameters cs and ca = 0). The analytic reading step proceeds in the same way as previously but with this isolated grapheme alone as focal. At this point, it should be noted that the isolated graphemes have been presented to the network in learning with their most frequent (unchronological) pronunciation (e.g., k → /k*/), where the silent phoneme /*/ plays the role of a syllable nucleus in phonological coding). As described in the next section, it was found in the simulations that graphemic-level processing occurred in exceptional situations, when the current focal window included an unusual although existing (e.g., #gal) or a nonexistent but pronounceable (e.g., #krou, #groie, #don) syllable. For example, the pseudoword krouba was read in the analytical mode in three consecutive steps: k → rou → ba → /k*/ (graphemic level) → /kui/ (syllabic level) → /ba/ (syllabic level), where only the first step was processed at the graphemic level. The syllable krou was not read as a whole because it includes the illegal although pronounceable orthographic sequence kr.

Finally in the analytic reading procedure, each successive phonological echo occupies only the clusters belonging to the first syllabic subset of the P layer, because the connections linking EM segment units to the first P subset alone will have been strengthened during learning. It is assumed that these successive segmental phonological echoes would be temporarily maintained and sequentially assembled in a phonological buffer (for neural network models of short-term storage of temporal sequences, see Ans, 1990b; Ans et al., 1994; Reiss & Taylor, 1991). The whole assembled phonology might then be transferred to the production system for global pronunciation of the input orthographic object.

**Permanent Learning**

It is noteworthy that the present network has the ability to learn what it reads on the assumption that it shifts to learning mode each time an orthographic input has just been read. When an input letter string has been read globally, the whole phonological echo that has just been computed remains self-maintained over the P layer (this property resulting precisely from the self-activation of the winning unit within a WTA cluster) while the global orthographic input remains available over the O1 layer. The network, which is now in a situation similar to that of learning in the global mode, can then learn the input letter string pronunciation.

When an input letter string is read analytically, the same principle applies for each segment (typically a syllable) that is successively read. In a given reading step, the phonological echo that occupies the first syllabic subset of the P layer is self-maintained. The focal part of the O1 orthographic input then consists of the segment that has been specified by the tagging signal from O2 to O1. The network, which would be at this time in a situation similar to learning in the analytic mode, can learn the pronunciation of each segment (within its orthographic context) of the input letter string.

It is further assumed that the network also has the potential ability to self-learn globally when the input letter string has just been read in analytic mode. On the one hand, the whole assembled phonology of the letter string is kept available in the phonological buffer at the end of analytic reading. On the other hand, the input letter string is simultaneously captured as a whole. It follows that, in these circumstances, the network will be in a situation similar to global learning.

In short, the network might create an EM word trace each time a word has just been read globally, whereas both a word trace and the corresponding segment traces might be created following analytic reading (with storage of their associated orthography-to-phonology correspondences).

**Word Frequency Account**

It has been stated that as many additional EM traces are created for the same orthographic event (a word or a segment) as have occurred in distinguishable states of the environmental context. The estimation of the number of traces created for a repeated orthographic event is the main object of the present section.

It is first assumed that the random environmental context associated with any orthographic event can take a number L of distinguishable states. Consider now the case where n EM traces have already been created for a given orthographic event (meaning that this same event previously occurred within n distinct environmental contexts). The probability q(n) that this orthographic event will occur within a new context (i.e., different from the n distinct contexts previously associated with this event), hence giving rise to an additional separate EM trace, is as follows:

\[ q(n) = \frac{[L - n]}{L} \]

Let \( \pi(z) \) be the mean number of EM traces created for \( z \) prior occurrences of a given orthographic event. Then, after a new occurrence of this same event

\[ \pi(z + 1) = \pi(z) + \frac{[L - \pi(z)]}{L} \]

It can be easily shown that the general term of this progression may be well approximated by the following function:

\[ \pi(z) = L(1 - e^{-\mu z}) \]

(5)

where it can be seen that the mean number of EM traces \( \pi(z) \), assigned to \( z \) occurrences of a given orthographic event, is bound by the number \( L \) of possible contextual states. Hence, the mean number of EM traces created for a given orthographic event is typically lower than its number of occurrences. The estimation of the mean number \( \pi(z) \) differs according to the nature of the repeated orthographic event, whole word or segment.

**Number of EM Word Traces**

With respect to words, let \( z_T(T) \) be the number of occurrences of a given printed word \( i \) up to some time \( T \) of learning. It is assumed that, in humans, this number increases with everyday reading experience although the relative word frequency remains constant over time. It then follows that the number of occurrences of any word \( i \) up to a given time \( T \) can be seen as
proportional to its frequency in the written language, noted \( f_i \), which gives

\[
z_i(T) = \lambda_T f_i \text{ for all } i,
\]

(6)

where \( \lambda_T \) is a factor that increases with experience.2 Substituting Equation 6 for \( z \) in Equation 5 makes the mean number \( n_i \) of EM word traces created for a given printed word \( i \), at time \( T \), to be an explicit function of its frequency \( f_i \):

\[
n_i(T, f_i) = L(1 - e^{-\alpha_T}) \text{ with } \alpha_T = \lambda_T / L,
\]

(7)

where \( \alpha_T \) is a parameter that characterizes reading experience.

Let \( \xi \) be the subset of the \( n \) EM word traces that were created by the multiple occurrences of the same word \( i \) up to \( T \). All these \( n \) EM word traces consist of the same orthographic part (the same stored word \( i \)) and of the same learning processing mode \( RM = G \). Consider now a reading test performed at time \( T \) (in the global reading mode \( RM = G \)) on a given orthographic object (word or pseudoword). As previously noted, a reading test is always viewed here as being performed in an unusual experimental situation in which the contribution of the environmental test context on the EM layer activity may be ignored. Consequently, each of the \( n_i \) word units belonging to the subset \( \xi \) have the same activation \( a_i \) (defined in Equation 3) and thus the same output activity (given in Equation 4). It is then formally equivalent to substitute the EM subset \( \xi \) with a single "macro" EM word unit \( i \) whose activation \( a_i \) remains the same as before (i.e., the same as that of one elementary word unit belonging to the subset \( \xi \)) but with an output activity \( s_i \) that is the sum of the \( n_i \) identical elementary output activities, which taking Equation 4 into account yields

\[
s_i = n_i a_i^\prime / \sum_{k=1}^{N} n_k a_k^\prime,
\]

(8)

where \( N \) is the number of macro EM word units assigned to the \( N \) distinct words learned up to time \( T \) (the EM word units alone are recruited in global reading mode as seen previously). Replacing in Equation 8 each \( n_k \) by its mean value \( \bar{n}_k(T, f_k) \) given in Equation 7, it follows that the output activity \( s_i \) of a formally equivalent macro word unit \( i \) will be

\[
s_i = \varphi_i a_i^\prime / \sum_{k=1}^{N} \varphi_k a_k^\prime \text{ with } \varphi_i = 1 - e^{-\alpha_T f_i}.
\]

(9)

In the subsequent simulations (isolated letter string reading), Equation 4 is replaced, in the case of the global reading mode, by Equation 9, which works for the single macro EM word unit \( i \) assigned to each distinct word \( i \). Thus, only one EM word trace is assigned to a given learned word \( i \) with its related global orthography-to-phonology correspondence and, therefore, with a single connection weight \( \varphi_i \) (or \( \omega_{ij} \)) between the macro EM word unit \( i \) and P units \( j \) (or O2 units \( j \)). In Equation 9, \( \varphi_i \) (0 < \( \varphi_i < 1 \)) accounts for the learned-word frequency \( f_i \) and \( a_i \) (0 \( \leq a_i \leq 1 \)) represents the orthographic similarity between a learned word and a tested orthographic object.

---

**Number of EM Segment Traces**

Consider now the estimation of the number of EM segment traces that have to be assigned to a given orthographic segment (typically, a syllable within its orthographic context) belonging to a given word repeatedly processed in the analytic mode. It has been noted previously that the network is potentially able to self-learn. It was reasoned that each time a word is read in the analytic mode, then an EM word trace is (probabilistically) created together with the corresponding EM segment traces. Hence, as long as a word is read analytically (because it cannot be read globally) both word traces and segment traces accumulate within the EM layer. It is then expected that after a limited number of analytic readings (thus creating a number, say \( m \), of EM word traces and \( m \) segment traces per word segment), this word would subsequently be read in the global mode. Consequently, the subsequent global readings of this word will no longer lead to the creation of additional EM segment traces, the number of segment traces remaining therefore limited to \( m \). Thus, if it is assumed that the minimal number \( m \) of EM word traces allowing word reading in the global mode is roughly the same for any word, then the same number \( m \) of EM segment traces can be roughly assigned to any word segment.

The same formal equivalence as the one previously made for word traces can be applied to segment traces. Each subset of \( m \) elementary EM segment units (units coding for the multiple occurrence of the same orthographic segment) can be substituted with a single macro EM segment unit whose output activity is expressed as in Equation 8 replacing \( n_i \) with \( m \):

\[
s_i = a_i^\prime / \sum_{k=1}^{M} a_k^\prime,
\]

(10)

where \( M \) is the number of macro segment units assigned to the \( N \) distinct segments6 learned up to time \( T \) (it can be seen that the output \( s_i \) of any macro EM segment unit \( i \) is independent of \( m \)). In the analytic reading mode, Equation 10 works for a single macro EM segment unit \( i \) assigned to each distinct orthographic segment \( i \); only one EM segment trace, with its related segmental orthography-to-phonology correspondence, is assigned to a given orthographic segment.

It is further noted that a formal EM macro unit represents a subset of elementary traces that are not assumed a priori to be spatially grouped like in topological maps (Ans, 1989; Kohonen, 1988). Rather, they are spread over the whole EM layer as in the ART model of Grossberg (1987a, 1987b). As a result, the episodic memory EM should therefore be resistant to local damage, although this property is not apparent when using the formal notion of a single macro unit associated with a given word or segment.

---

6 It was previously shown that the reading network has the ability to learn what it reads. This permanent learning ability makes it possible to consider that the network has been exposed to a number of word occurrences similar to that of individuals.

6 Recall that any orthographic word segment is always considered within its orthographic context. For example, \( \text{cho} \) in \( \text{chocolat} \) and \( \text{cho} \) in \( \text{chorale} \) are two distinct orthographic segments.
Simulations

The Training Word Base and Model Parameters

The reading network works on a training base of 13,165 mono- and polysyllabic French words that were extracted from Brulex, a database of 35,746 lexical entries (Content, Mousty, & Radeau, 1990). All redundant entries corresponding to the same word coded with a different grammatical class were first eliminated from the original database in order to keep only one occurrence per word. Most words of very low frequency (<100 occurrences per 100 million) were also deleted as these words were very often unknown to even very skilled readers. However, some low frequency irregular words that were likely to belong to the vocabulary of an average reader were retained along with their matched regular controls. Most words for which the lexical frequency was missing in the original database were further eliminated. These words were typically the plural (e.g., garnis—lined) or feminine (e.g., garnie) forms of words appearing in the database under their singular masculine form (e.g., garni). Nevertheless, those words for which the orthography or pronunciation differed for the masculine and feminine forms (e.g., acteur—actor and actrice—actress) or for the masculine and plural (e.g., cheval—horse and chevaux—horses) were retained; the frequency of the singular masculine form was by default assigned to these words. Finally, a few polysyllabic words having highly irregular orthographic and phonological characteristics (e.g., those taken from foreign languages, e.g., adagio /adadjə/) were also eliminated because they could not be automatically segmented into syllables.

The remaining 13,165 words (from 1 to 10 letters and one to five syllables) had an orthographic frequency (Imbs, 1971) between 4 and 5,607,822 occurrences per 100 million (median frequency = 553). The main characteristics of this set of words are given in Table 1.

The phonological codes assigned to the items in the training base used in the present model were generally equivalent to the phonetic codes of Brulex because the phonetic form of an isolated word directly derives from its phonological specification. However, the phonetic codes of Brulex that correspond to the Parisian French dialect were modified for some words. Indeed, in Parisian French, most words with a grapheme e (e.g., souvenir, cheval, petit, table, or forte) are pronounced in their shortened form (/suvniR/, /Sval/, /pti/, /tabl/, or /foRt/) when it is typically assumed that these words include, at a deep level, a vowel schwa (/suvniR/, /Sval/, /p̠ti/, /tabl/, or /foRt/) that can be phonetically realized or not (Charette, 1991; Dell, 1973; Encrevé, 1988; Selkirk, 1978; Tranel, 1987). The phonetic realization of schwas is determined by various factors. It is subject to dialectal variations: Most schwas are phonetically realized in some dialects (e.g., in southern French: Durand, Slater, & Wise, 1987), whereas they are typically deleted in other dialects (e.g., Parisian or Northern dialects). The phonetic realization of schwas is also modulated, for the same word and the same speaker, by speech speed and pragmatic constraints; schwas are more frequently realized phonetically when speaking slowly, in situations of environmental noise, or when speaking to a child or a foreigner. Finally, schwas are always represented in spelling by the letter e, and they are more systematically realized phonetically in reading than in speaking, at least by beginning or unskilled readers. The output of the model being a phonological representation, a vowel schwa was systematically reintroduced in words for which the phonetic representation of Brulex (at a surface level) was shortened. This was done in order to have a more conservative phonological representation (at a deeper level) that could ultimately be realized through a variety of phonetic forms.10 The list of phonemic notations adopted in the simulations is given in Table 2.

The 13,165 words in the training base were presented to the model; polysyllabic words were learned in both the global and the analytic mode, whereas monosyllabic words were only learned in the global mode. As seen in the previous sections, learning these words in global mode gives rise to \( N = 13,165 \) distinct macro EM word traces (with their \( N \) related global orthography-to-phonology correspondences stored in the network connections). Factor \( \phi_i \), which takes into account word frequency \( (f_i) \), was assigned to each EM word trace, which was further marked with its associated processing mode \( RM = G = 10 \). Learning polysyllabic words in the analytic mode gives rise to 35,822 EM syllabic segment traces and an equal number of syllabic orthography-to-phonology correspondences. The EM layer also contains 75 (graphemic) segment traces that have

---

10 Consider, for example, the French word maintenant (now), for which the phonetic form assigned in Brulex is /mɛnən/. In fact, this word can be produced under a number of different forms in conversational speech, being pronounced either /mɛnən/, with a preserved schwa, or /mɛnə/, after schwa deletion, or /mɛn/, the deletion of schwa leading to the /t/ deletion. The more conservative form /mɛn/ has been here assigned to this word.

---

Table 1

<table>
<thead>
<tr>
<th>Characteristics of the Word Base</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of letters</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Word length</td>
</tr>
<tr>
<td>1–2</td>
</tr>
<tr>
<td>3–4</td>
</tr>
<tr>
<td>5–6</td>
</tr>
<tr>
<td>7–8</td>
</tr>
<tr>
<td>9–10</td>
</tr>
<tr>
<td>Word frequency (per 100 million)</td>
</tr>
<tr>
<td>Range</td>
</tr>
<tr>
<td>4–50</td>
</tr>
<tr>
<td>50–100</td>
</tr>
<tr>
<td>100–150</td>
</tr>
<tr>
<td>150–200</td>
</tr>
<tr>
<td>200–300</td>
</tr>
<tr>
<td>300–400</td>
</tr>
<tr>
<td>400–600</td>
</tr>
<tr>
<td>600–1,000</td>
</tr>
<tr>
<td>1,000–2,000</td>
</tr>
<tr>
<td>2,000–5,000</td>
</tr>
<tr>
<td>5,000–10^4</td>
</tr>
<tr>
<td>10^4–5.10^4</td>
</tr>
<tr>
<td>5.10^4–10^5</td>
</tr>
<tr>
<td>10^5–6.10^6</td>
</tr>
</tbody>
</table>
been created when learning the pronunciation of some isolated consonantal and vocalic graphemes (see the list in Appendix A); this gives a total of $M = 35,897$ EM segment traces that have no associated frequency factor (as seen previously) but are marked with $RM = A = 0\{1\}$, orthogonal to $G$. Setting up orthography-to-phonology correspondences in the analytic reading mode further required the prior development of segmentation algorithms. These algorithms were used for segmenting the phonemic codes of polysyllabic words into syllables and each syllable into onset and rime; they also automatically segmented the orthographic codes of these words in a similar way.

Any reading test of an orthographic object (word or pseudoword) is always initiated in the global reading mode $RM = G$. In this mode, the parameter that modulates the partial contrast enhancement within the EM layer is set to $\gamma = 13$; the thresholds of the orthographic O2 units and of the phonological P units are set to $\theta_O = 0.35$ and $\theta_P = 0.17$, respectively. When the system runs in the analytic reading mode, these values become, respectively, $RM = A$, $\gamma = 35$, $\theta_O = \theta_P = 0.1$. With respect to the syllabic level processing, the context size parameter is set to $cs = 1$ (only one contextual letter) and the O1 units coding for contextual characters have an activity equal to $ca = 0.75$. In the graphemic-level processing, the orthographic context is not taken into account, so that $cs$ and $ca = 0$. The word frequency factor $\varphi_i = \varphi_i(f_i)$, which, expressed as Equation 9, is determined by the value taken by the parameter $\alpha_T$ (defined in Equation 7) at a given time $T$ (when the reading tests were performed). Except in the Developmental Perspective section, where this parameter value was varied, all of the following reading simulations were performed using the fixed value $\alpha_T = 10^{-3}$. The choice of this $\alpha_T$ value was directed by simple heuristics. The histogram of the $\varphi$ values is relatively uniform for $\alpha_T = 10^{-3} (M = 0.508, \text{Mdn} = 0.424, \text{range} = 4.10^{-3} - 1)$, whereas it is skewed for $\alpha_T = 10^{-4}$ (shifted towards the low $\varphi$ values, $M = 0.157$, $\text{Mdn} = 0.053$, range $= 4.10^{-2} - 1$) and for $\alpha_T = 10^{-5}$ (shifted towards the high $\varphi$ values, $M = 0.929$, $\text{Mdn} = 0.996$, range $= 4.10^{-4} - 1$). Furthermore, if one considers the median frequency value ($f_{\text{med}} = 600$) and the highest frequency value ($f_{\text{max}} = 6 \times 10^6$) of the database, their ratio is approximately equal to $10^4$. The $\varphi$ ratio corresponding to these two frequencies ($\varphi_{\text{max}}/\varphi_{\text{med}} = 2.2$) is equivalent to the ratio that would be obtained using a classical frequency logarithm ($\log \varphi_{\text{max}}/\log \varphi_{\text{med}} = 2.4$), whereas the $\varphi$ ratios for $\alpha_T = 10^{-4}$ and $\alpha_T = 10^{-2}$ are near to 17 and 1, respectively.

Word Reading

The 13,165 words learned by the network were presented for reading. Most words ($n = 12,839$, or 97.52%) were processed globally (reading mode $RM = G$); the remaining words ($n = 326$, or 2.48%) were processed in the analytic mode ($RM = A$). Analytic processing essentially applies to words of the lowest frequency ($\text{Mdn} = 131$, range $= 4\ldots536$), the probability that a word will be processed globally increasing with its frequency. This is shown in Figure 6.

A pronunciation was accepted as correct when the output of the network was strictly identical to the target (i.e., contained all target phonemes in their expected position). The network generated the correct phonology of most previously learned words (96.25%). The number and percentage of words correctly read in each reading mode are given in Table 3.

Error Analysis

The network produced 3.68% of erroneous outputs ($n = 472$) in the global mode. Errors typically concerned low-frequency words, the median frequency of erroneous responses being 204 (range $= 29\ldots45,134$) against 612 (range $= 17\ldots5,607,822$) for the correct productions. More than half of the errors (57.63%) consisted of the reduplication of one target phoneme (e.g., /croisi/ $\rightarrow$ /koRROzi/ instead of /koROzi/, /chariot/ $\rightarrow$ /SaRRIO/ instead of /SaRIO/); the output then included all the expected phonemes in the expected positions plus one duplicated phoneme. This kind of phoneme duplication occurred only at syllable boundary and was therefore found on polysyllabic words. Errors of addition were also produced (e.g., /dudative/ $\rightarrow$ /dükaküvi/ instead of /dükaktiv/). In this latter case, another real word differing by a single phoneme from the target was produced instead of the input word in 10.16% ($n = 48$) of the erroneous productions. In most cases ($n = 45$), the production was of higher frequency than the target (see the word list in Appendix B). An additional consonantal phoneme was also sometimes ($n = 25$, or 5.3%) introduced at the beginning of
words with a vocalic onset (V → CV; e.g., *incisive* → /désiziv/ instead of /ésiziv/) or between two vocalic phonemes (VV → VCV; e.g., *écuelle* → /ékylœl/ instead of /ékylœl/ or *poète* → /poçêt/ instead of /poçêt/), thus leading to a sequence characterized by a simpler syllabic structure. Occasionally (n = 10), the error concerned a low-frequency irregular word that was then regularized (e.g., *pays* /pêl/ → /pai̞/; *broc* /bRO/ → /bROkl/, *pouls* /pu/ → /pul/).

Twenty-one errors were collected in the analytic mode. Six errors consisted of regularizations of low-frequency irregular words (*suspense* /sySpens/ → /syspêns/; *revolver* /Révolvér/ → /R′ yolvér/, *gourd* /guR/ → /guRd, kaiser/ /kézêR/ → /kézê/, *quartz* /klaRts/ → /kaRtz/, *boycott* /bOIKotkl/ → /bûak otkl/). The other erroneous productions were due to confusions between /e/, /ê/ and /ë/ (e.g., *feuillé* /fejê/ → /fêjê/ or *clocheton* /kloSh ê/ → /kloSêto/).

**Pseudoword Reading**

A list of 830 mono- and polysyllabic pseudowords was designed. Pseudowords (from one to three syllables and from three to eight letters) were made from real words by changing one letter or more while preserving their length and syllabic structure. For example, the pseudowords *fiarté, clirat, bourgon, rulleetin,* and *glaud* differ by one letter (or one phoneme) from the real words *clarte, climat, fourgon, bulletin,* and *gland,* respectively, whereas the pseudowords *plordé, dricat, veurdin, rallition,* and *bront* differ from the same words by more than a single letter (or phoneme).

When the 830 pseudowords were presented to the model, 22.65% were processed globally (n = 188), although most were read analytically (n = 642, 77.35%). In RM = A, 529 items (82.40%) were processed syllable by syllable and 113 items (17.60%) at the graphemic level. This latter level of processing typically involved a single grapheme (in fact, a single letter) of the string (n = 100, 88.50%), the remaining being processed at the syllabic level; in only one case (for the pseudoword *groie*) were two graphemes of a pseudoword processed at the graphemic level. It is noteworthy that the isolated grapheme traces virtually never contribute to the pronunciation of the input letter strings when they are processed at the syllabic level. In order to determine whether the output of the network corresponded to "correct" pseudoword pronunciation, 20 participants were presented with the 830 pseudowords to read aloud.

Their responses were recorded and compared with those of the network. In the network, a correct response to a pseudoword was defined as a phonologically plausible response that was produced by any 1 of the 20 participants. Overall, the network correctly pronounced 89.76% (n = 745) of the pseudowords (see Table 4), typically (in 99% of cases) generating the pronunciation most often attributed to pseudowords by human participants.

As shown in Table 4, errors were more likely to occur when the pseudowords were processed globally. In this case, errors frequently concerned pseudowords that differed from a real word by only one letter, many erroneous responses (n = 31 of 51, or 60.78%) being lexicalizations in this case. The list of the 31 lexicalizations is given in Appendix C (a comparison with the performance of human participants is presented below). The other 20 errors were productions that differed by one phoneme (which was either substituted, omitted, added, or duplicated at syllabic boundaries) to the target (e.g., *gridon* /gRidô/ → /pRidô/; *sécot* /sêko/ → /sêo/; *phage* /faj/ → /pfaj/; *calier* /kalï/ → /kalï/).

Half of the 34 errors collected in the analytic reading mode concerned pseudowords with glides. Only two of these errors were observed following processing at the grapheme level (lions /lwos/ → /lwos/ and *siomte* /s ôt/ → /sîmô/). The other half were errors of reduplication observed on polysyllabic items at syllabic boundaries.

**N Count and Length Effects**

As shown in Figure 7, the probability that a pseudoword would be processed globally increased with its number of neighbors, whereas the probability that it would be accurately read in the global mode remained roughly constant whatever its N count. The global processing of pseudowords also varied as a
function of word length. Figure 8 clearly shows that the probability that a pseudoword would be processed globally decreased as its length increased. In fact, length and N count were negatively correlated ($\rho = - .587$), so that the length effect could be secondary to the N-count effect. This seems to be the case, because an identical N-count effect was observed on those pseudowords of constant length.

In sum, the network reads both polysyllabic words and pronounceable polysyllabic nonwords with high accuracy because 96.25% of words and 89.76% of pseudowords were correctly pronounced. The model’s performance therefore demonstrates that a good level of accuracy in pseudoword reading can be achieved without resorting to an explicit GPC rule system as previously shown by Plaut and McClelland (1993) and PMSP96, contrary to the claims of dual-route theories of reading (Besner et al., 1990; Coltheart et al., 1993). In the present network, however, words were typically processed globally and pseudowords analytically; this result is more in accordance with dual-route models but contrasts with PDP models in which processing always operates over the entire letter string. It is noteworthy, however, that none of the present model’s reading modes is exclusively dedicated to a specific category of stimuli because a number of pseudowords were successfully processed globally, that is, on the basis of word traces alone (see Appendix D for an example showing the generalization process based on the integration of similarities between an input pseudoword and learned whole words), whereas a number of words, including some irregular words, were accurately pronounced analytically.

### Comparison of Simulated Results With Human Performance

**Lexicalization Errors in Pseudoword Reading**

As shown in Table 4, the system generated 51 erroneous responses when reading 188 pseudowords in the global mode. Of these 51 errors, 31 were lexicalizations. The lexicalization errors in fact occurred on 17.8% (31 of 174) of the pseudowords with lexical neighbors, but none occurred on pseudowords without neighbors. Other error types were also observed; they concerned 8% (14 of 174) of the pseudowords with neighbors and 42.8% (6 of 14) of the other pseudowords. Overall, 68.9% of the observed errors on pseudowords with neighbors were lexicalizations. We attempted to determine whether human readers would also show a similar tendency towards lexicalization in experimental conditions likely to favor global processing. In order to elicit global processing in human participants, the list included a majority of real words mixed with pseudowords so that participants were presented with stimuli for which lexicality was not predictable (mixed presentation) and for which successive encoding of visual information was limited (brief exposure). For this purpose, 40 participants were asked to read aloud a list of 60 regular high-frequency words and 40 pseudowords for which 20 had lexical neighbors and 20 did not. Items were from five to eight letters and one to three syllables in length. The words and pseudowords were mixed and presented in a random order for only 83 ms (followed by a mask) on a computer screen. In this experiment, the participants’ reading performance was rather poor on pseudowords (60.5% correct on average) but almost perfect on words (98.9% correct). With respect to the error types, on average 35.5% (ranging from 15% to 60%) of the pseudowords with neighbors were lexicalized, whereas lexicalizations concerned only 1.1% of the pseudowords without neighbors. In fact, 93% erroneous responses were lexicalizations in the former case but only 2% in the latter. These results therefore demonstrate that human participants exhibit a strong tendency towards lexicalizations in pseudoword reading when placed in experimental conditions that limit the use of analytic strategies (i.e., brief stimulus presentation).

Furthermore, it is noteworthy that although the participants’ ability to read pseudowords was weakened (60.5% correct) when they were placed in conditions that limit analytic processing, their performance in reading real words (under the same conditions) remained unchanged (98.9%). This is consistent with our assumptions that two procedures underlie skilled reading.

**Word and Pseudoword Naming**

A comparison with human performance was conducted in order to verify that the network’s general level of accuracy on words and pseudowords would be similar to that of skilled readers. Eighteen human participants were asked to read as quickly and accurately as possible lists of 80 words and 80 pseudowords (from the original list of 830 pseudowords) that were strictly matched for length (five to six letters and one to three syllables), neighborhood size, and phonological and orthographic structure (Bozon & Carbonnel, 1996). Items were mixed and presented in a random order on a computer screen. The mean reading performance of the human participants was compared with that of the network when processing the same lists of items. In order to verify the stability of the network performance, it was further tested on an extended list of 512 words and 268 pseudowords having the same characteristics of length, frequency, and N count as the experimental items. These additional items were also extracted from the original training list of 13,165 words and 830 pseudowords. The results are summarized in Table 5. They show that the network’s performance accurately simulates human performance.
List Effects on Pseudoword Naming

The results that were obtained in mixed presentation of words and pseudowords were comparable to the overall results of the network (summation of performance obtained in analytic and global reading mode). Indeed, in these experimental conditions, participants performed at a level (92.43% correct on pseudowords; see Table 5) quite similar to the overall performance of the network on the same experimental list (90%) as well as on a matched extended list (88.81%). The overall performance of the network on the initial list of 830 pseudowords was also very close to these results with 89.76% correct responses (see Table 4). This level of performance was compared to that of the 20 participants previously asked to name the list of 830 pseudowords in blocked presentation conditions. In this particular experiment, participants were informed about the nonlexical nature of the items to be read. Their responses were analyzed, counting as incorrect those responses that included either a substitution (e.g., "/vanici/ /vanisi/ /vanisi/"), an addition (e.g., "/soubier/ /subi/ /suRbi/"), or a transposition (e.g., "/tapRi/ /tRapii/"). On average, human participants accurately read 814 of 830 pseudowords (98.1%) performing this task at a level higher than the overall performance of the network.

It can be assumed that in conditions in which the nonlexical nature of items is known a priori, human participants are less likely to use an orthographic check. It follows that in these particular experimental conditions (blocked lists of pseudowords, a great number of items), human participants’ performance would be more appropriately compared with that of the network when running in analytic mode. It has already been observed (see Table 4) that the network, when processing analytically, accurately read 94.70% of those nonwords (n = 642). A similar performance (94.46%) was obtained on the whole list of 830 pseudowords when the network was forced to read them all analytically. Therefore, and as expected, the performance of the system in analytic mode paralleled that of human participants when reading the 830 pseudowords under conditions of blocked presentation.

Neighborhood Effects in Perceptual Identification

It has been found that neighborhood effects do not manifest similarly in tasks involving the recognition of an orthographic string, such as in lexical-decision or perceptual identification tasks, and in tasks requiring the production of a phonological code, such as naming (Grainger, 1990, 1992). The neighborhood effects related to consistency that are observed in naming are investigated in the next section. The performance of the network here is compared with the performance of skilled readers when engaged in a task of letter-string identification under conditions of very brief stimulus exposure (Bozon & Carbonnel, 1996). This simulation was done with the purpose of simulating the striking lexicality by neighborhood interaction that is observed in this task in human participants. The experiment consisted of presenting words and pseudowords on a computer screen for 33 ms followed immediately by a pattern mask. Participants were asked to attend to the target and write down what they had detected immediately after presentation.

The rationale for simulating this experiment is as follows. It is assumed that under conditions of very brief exposure, the orthographic input over O1 will have decayed at the time the orthographic echo is established over O2. As a consequence, a comparison between the orthographic input on O1 and the orthographic echo generated over O2 will be prevented, so that the system becomes unable to shift into the analytic mode. The orthographic echo generated by the system when clamped in global reading mode should then grossly correspond to the written performance of human participants. This hypothesis was verified by comparing the participants’ performance on a list of 80 words and 80 pseudowords varying on N count with the model’s O2 outputs obtained on the same lists of words.

As shown in Table 6, the participants’ results in the word identification task demonstrate a negative neighborhood effect that is only significant for low-frequency words. In contrast, a positive neighborhood effect characterizes pseudoword processing. The same pattern of results characterizes the performance of the network with the absence of neighborhood effects.
on high-frequency words, a negative effect on low-frequency words, and a positive one on pseudowords. Essentially similar results were also obtained on extended lists of 512 matched words and 268 pseudowords.

A detailed post hoc analysis of the participants’ responses was further conducted and compared with the network’s orthographic outputs in order to determine whether the performance of the system was qualitatively similar to that of the human participants. The analysis was restricted to the pseudowords, the error rate on words being too small to allow comparison. The rate of correct identification was then calculated for each letter in each position of the output. The results from the analysis of the human data show that letters that belong to the first vocalic grapheme (FVG) of pseudowords were far more likely to be correctly reported than any other letter of the string. Indeed, the FVG letters were correctly reported at a rate of 86% against only 67.5% for the other letters (non-FVG letters). Moreover, this difference cannot be attributed to a simple position effect; although the FVG letters more frequently occurred in the second and third positions of the experimental pseudowords, the non-FVG letters when occupying the same second and third positions were only identified at a rate of 69%. The difference cannot be attributed to a simple consonant—vowel distinction; indeed, the FVG was on average more accurately identified (81.1% correct identification) than the non-FVG vowels (66.05%) in bisyllabic items. In the simulations, similar results were obtained on the experimental pseudoword list with a recognition rate of 94.17% on the FVG letters, 76.97% on the non-FVG letters, and 75.44% on those non-FVG letters occupying the target positions. Similarly, results were 97.42%, 78.16%, and 78.29% correct, respectively, on the extended pseudoword list. On both lists, the network performance was superior for the FVG letters as compared to the non-FVG vowels (97% vs. 78.79% on the experimental list and 98% vs. 83% on the extended list). The superiority of recognition of the FVG letters in the simulations straightforwardly follows from the adopted basic assumption that the system focuses on the FVG of the input string. Empirical data suggest that the same focalization process would also characterize reading in human participants.

### Naming Latency

So far, we have shown that the behavior of the network simulates human performance with respect to accuracy of response. It is also to be shown that the performance of the system parallels human data with respect to naming latency. The model postulates the existence of two reading modes with global processing applying first and analytic processing only proceeding when global processing has failed. An indicator of naming latency in global mode is here proposed and compared with experimental data. The evaluation of this source of naming latency is crucial for the simulation of the basic effects of frequency and regularity that concern real words (that are typically read globally by the network). On the other hand, any attempt to simulate latency in analytic mode would require a precise computation of the global-to-analytic shifting time as well as an evaluation of the time taken for the successive visual captures within the analytic mode itself. Such shifting operations have not been implemented; therefore, a true simulation of latencies in analytic mode is beyond the scope of the present article. However, some predictions can already be made on the difference of reaction times between items processed globally versus analytically.

It can indeed be assumed that naming latency should be longer for analytically than for globally processed printed stimuli because the analytic mode only applies after the global mode has failed. On the one hand, and as previously shown, virtually all words (97.5%) are processed globally, whatever their length. On the other hand, most pseudowords (in particular those without neighbors) are processed analytically. Consequently, the model predicts that most words, whether mono- or polysyllabic, should be named with shorter latencies than any pseudoword. In addition, an increase in reaction times with syllabic length is also predicted for pseudowords because analytic processing is sequential, each new syllable requiring a new visual capture of information. The experimental data on naming latency reported by Carbonnel et al. (1998) are in agreement with these predictions. The authors asked 32 skilled readers to name 60 words ranging from low to medium frequency (mean logarithmic fre-
frequency = 2.7) from one to three syllables (3 × 20). Words were matched in frequency over the three length categories. Participants also had to name 60 pseudowords (without neighbors) matched to words in length (number of syllables, graphemes, and phonemes), syllable frequency, and bigram frequency. Words and pseudowords were presented by blocks mixed for length within each lexical category. For each category of items (words and pseudowords of every syllabic length), naming latencies for each participant were arranged in ascending order and deciles were calculated. Deciles were averaged over participants to give group deciles following the Vincent (1912) averaging method. The group latency distribution histograms were then constructed (see Ratcliff, 1979) for each item category by plotting deciles on the abscissa and then constructing rectangles between adjacent deciles such that all the rectangles had equal areas (Figure 9). Thus for each histogram, the y-coordinate in some way reflects the probability density as a function of latency.

It can be seen in Figure 9 that for each syllabic length there is only a little overlap between words and pseudowords and that the distributions are all the more disjoint as syllabic length increases. Furthermore, the distributions of latencies for mono-, bi-, and tri-syllabic words are almost superimposed, whereas the corresponding distributions for pseudowords are largely disjoint. Finally, and as predicted, it can be concluded from the probability density distributions (in Figure 9) that any of the pseudowords are likely to be named with a longer latency than any of the words.

With respect to global processing, the proposed account of latency is as follows. The phonological layer contains one cluster for each potential phonemic position. Each cluster runs in a WTA manner, the winner unit (if at all) taking a final activity value of 1 and the other units 0; the phonological response of the network is therefore explicit and clean (unequivocal). It can therefore be assumed that the source of latency variation lies in the clean-up process within WTA clusters. The rationale for estimating the time needed for cleanup (denoted tcl) within a given cluster is as follows.

Let p1 and p2 be the above-threshold input activation (as previously defined in the Model section) of the winner unit within a P cluster and its main competitor (if any), respectively, and δ = p1 − p2, a value reflecting the competition strength; the clean-up time tcl will be a decreasing function of p1 when no competition occurs within the cluster and of both p1 and δ when a competition does occur. A simple way to express this function is

\[ t_{cl} = B[\log_{10}(1/p1) + \log_{10}(1/\delta^c)], \]

where B is a constant (B = 100 in all subsequent simulations), c = 0 in the absence of competition, and c = 1 otherwise. This expression can be rewritten as

\[ t_{cl} = B \log_{10}[1/(p1.\delta^c)]. \]  

Figure 10 illustrates how the clean-up time tcl varies as a function of p1 and δ values. Notice that competition, when it is strong (weak δ values), disproportionately influences the tcl magnitude in comparison to those cases without competition.

Let us now consider the case of a word read in global mode; activation from layer EM is transmitted in parallel to all positional P clusters and these clusters also stabilize in parallel. At this point, we make the assumption that the phonological response of the network has to be entirely stabilized (i.e., all the P clusters have to be clean) before transmission to the articulatory system. This implies that the cluster that determines the latency (denoted Lat) of the system for a given word read globally is the cluster that has the longest clean-up time. It follows that

\[ \text{Lat} = \max_k \{t_{cl}\}, \]

with k ranging over all P clusters.

**Frequency and regularity effects.** A simulation of the regularity effect in the present network obviously requires reference to experimental data relative to the French language, a language that differs from English with respect to the notion of irregularity on three main aspects. Irregularities are less numerous in French than in English, they frequently concern consonants or consonant clusters rather than vowels, and they are more frequently observed within low-frequency words. It further appears more appropriate to simulate results from an experiment using mono- and polysyllabic items because one of the main feature of the present network is its ability to process polysyllabic items. In French, some irregular polysyllabic words have enemies (i.e., neighbors that share the irregular spelling unit with the target but not its phonological counterpart—e.g., irregular word: laser /lazeR/; enemy: lever /lave/), as is the case for most monosyllabic irregular words in English, but a good number of irregular polysyllabic words have no enemy at all. For example, the word orchidée /orKide/, although having no enemy neighbors, is typically considered as being irregular because it includes a spelling pattern, 'ch', that is pronounced /S/ in most mono- and polysyllabic French words (e.g., in chic, archive, chimère, or machine).

The use of polysyllabic items therefore provides the opportunity to dissociate the effects of regularity and neighborhood consistency on naming latency. Peereman (1995) designed an experiment on French polysyllabic words investigating the performance of skilled readers in pronouncing regular and irregular words of high and low frequency. His study further addressed the role of neighborhood consistency in naming low-frequency exception words. The notion of neighborhood was extended so that not only words of similar length but also words longer or
shorter by one letter than the target word were considered as its neighbors. The presence or absence of enemies and their relative frequency were manipulated for low-frequency words only. Peereman’s adult participants performed as follows. Low-frequency exception words with frequent enemies were pronounced significantly more slowly ($M = 547$ ms) than their regular controls ($M = 490$ ms). However, this classical regularity effect was not observed for low-frequency exception words without enemies; a nonsignificant trend for naming these items with a shorter latency ($M = 501$ ms) compared with their regular controls ($M = 516$ ms) was even reported. A difference of 28 ms was further obtained between exception and regular high-frequency words that was significant in the subject analysis but not in the item analysis. Overall, the most striking result of Peereman’s study was the demonstration that the effect of regularity is not systematically observed within low-frequency words but depends on neighborhood consistency.

This demonstration of a reliable interaction between regularity and neighborhood consistency among low-frequency words is a challenge to any model of skilled reading. The DRC version of the dual-route model (Coltheart et al., 1993; Coltheart & Rastle, 1994) predicts a regularity effect on low-frequency words because of a conflict between outputs from the lexical and nonlexical procedures at the phoneme stage. A stronger regularity effect might then be expected for low-frequency words with frequent enemies compared with low-frequency words without enemies if one admits that the concurrent phoneme receives activation not only from the nonlexical procedure but also from the lexical procedure when neighbors that disagree with the target do exist. In any case, the model predicts longer naming latency for all low-frequency irregular words (irrespective of neighborhood consistency) relative to their regular controls because some conflicting information in all cases arises from the nonlexical procedure. Obviously, such a prediction is in disagreement with Peereman’s (1995) finding that low-frequency exception words without enemies do not yield longer naming latencies than low-frequency regular words.

The ability of the present network to simulate the regularity

---

**Figure 9.** Distribution of naming latencies for words and pseudowords according to syllabic length.

**Figure 10.** The clean-up time $t_{cl}$ within a phonological cluster as a function, expressed in Equation 11, of $p_1$ and $\delta = p_1 - p_2$, where $p_1$ and $p_2$ are the above-threshold input activations of the winner unit and its main competitor and $\delta$ a value reflecting the competition strength within the cluster. Only the range of variation of $(p_1, \delta)$ allowed by the model’s constraints is considered: $p_1, p_2 \approx \theta_p = 0.17$ and $p_1 + p_2 \leq 1$ (because the total activity of episodic memory layer EM equals 1). NC = no competition.
by neighborhood consistency interaction was assessed using Peereman’s (1995) lists of unfamiliar exception words with frequent enemies (mean frequency = 339, SD = 60) and their matched regular words (mean frequency = 328, SD = 64) and unfamiliar exception words without enemies (mean frequency = 341, SD = 81) and their matched controls (mean frequency = 356, SD = 65). The experiment with familiar regular words (mean frequency = 40,165, range = 5,666–129,650) and familiar exception words (mean frequency = 45,928, range = 6,258–417,000) was also simulated.

The performance of the network (see Table 7) reveals that low-frequency exception words with frequent enemies show a higher mean latency score than their matched regular control words. In contrast, the mean latency obtained for low-frequency exception words without enemies is even (although not significantly) lower than that of their regular controls. The performance of the model is therefore consistent with Peereman’s (1995) results. Moreover, when the regularity effect (for both high-frequency and low-frequency words) is evaluated through the difference between the mean latency scores in each list, the regularity pattern obtained in the simulations strikingly parallels that reported in human data as shown in Figure 11.14

Peereman’s (1995) results further suggest that the often emphasized Regularity × Frequency interaction should in fact be expected only when low-frequency exception words have frequent enemies in their neighborhood. In the absence of empirical data addressing the effect of neighborhood consistency for high-frequency words, a simulation was conducted using only the lists of low-frequency exception words with and without enemies and their matched regular controls. Frequency was artificially manipulated so that each word appeared one time in the low-frequency list with the same (low) frequency as in the previous simulation and one time in the high-frequency list with a pseudo high frequency.15 Figure 12 shows the results of the simulation; frequency interacts with regularity only when exception words have high-frequency enemies, F(1, 22) = 6.37, MSE = 1,169, p = .019, for exception words with high-frequency enemies (E+); F(1, 23) = 1.67, MSE = 162, p = .21, for exception words without enemies (EO). The effect of frequency is significant for both E+ and EO, F(1, 22) = 24.06, MSE = 1,169, p = .0001, and F(1, 23) = 14.92, MSE = 162, p = .0008, respectively. There is a marginal main effect of regularity only for E+, F(1, 22) = 4.44, MSE = 2,753, p = .047. Furthermore, pairwise comparisons reveal that the regularity effect is significant on low-frequency E+ words, F(1, 38) = 9.88, MSE = 1,961, p = .003, but not on high-frequency E+ words, F(1, 38) = .15, MSE = 1,961, p = .7.

Thus, the Regularity × Frequency interaction typically re-

<table>
<thead>
<tr>
<th>Word category</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low frequency</td>
<td></td>
</tr>
<tr>
<td>Exception: frequent enemies</td>
<td>92.4</td>
</tr>
<tr>
<td>Regular controls</td>
<td>35.4</td>
</tr>
<tr>
<td>Exception: no enemies</td>
<td>12.6</td>
</tr>
<tr>
<td>Regular controls</td>
<td>24.5</td>
</tr>
<tr>
<td>High frequency</td>
<td></td>
</tr>
<tr>
<td>Exception</td>
<td>17.1</td>
</tr>
<tr>
<td>Regular controls</td>
<td>8.9</td>
</tr>
</tbody>
</table>

Figure 11. The size of regularity effect as measured by naming latency according to word type: latency difference, in an experiment with human participants (adapted from Peereman, 1995) and in the network simulation, between a list of exception words and a list of regular controls (the same lists in experiment and simulation), for low-frequency (LF) exception words when they have frequent enemies (LF E+) or no enemies (LF EO) and for high-frequency words (HF).

14 The simulated and experimental latency patterns can only be compared here if it is assumed that both early perceptual processing and motor planning introduce only minimal latency variations from one list to the other. In Peereman’s experiment, the difference of latency between immediate and delayed naming remains almost constant (around 350 ms) for all word types, so that latency variation in human performance seems essentially determined by phonological processing. Latency in the model is precisely determined at this same level.

15 A high frequency of 40,000 was attributed to each word in order to be comparable to the mean frequency of Peereman’s high-frequency lists.
words, a result not replicated in the last implementation of PMSP96, in which high-frequency irregular words were higher in latency than low-frequency regular words (0.034 vs. 0.014 as compared with a regularity effect for low-frequency words of 0.039).

The ability of the present model to account for these regularity effects straightforwardly follows from its basic principles. The activation of the phonological layer P results from the summation of all the global phonological correspondences set during learning, each contributing in proportion to the output value of the EM unit supporting its specific word trace—this value itself being a function of both the contrasted similarity of the orthographic input with the word trace and its associated frequency (see Equation 9). As classically defined, a word is considered irregular when it includes a grapheme that is associated with a concurrent (regular) pronunciation in most other words (referred to below as antagonistic words). When an irregular word is presented to the network, only those antagonistic EM word traces that are orthographically similar to the target and of high frequency, that is, those corresponding in fact to high-frequency enemies, are sufficiently activated to contribute significantly to phonemic activation within the P clusters. In this case, there is a strong probability that the activity of the target (irregular) phoneme will be relatively weak, that a (regular) phoneme competitor will emerge within the relevant positional P cluster, or both. In contrast, all the antagonistic word traces remain weakly activated when they share only a few orthographic features with the target or are of low frequency. In this latter case, they contribute only minimally, if at all, to phonemic activation, so that activity of the target phoneme is not expected to be hindered by the concurrent phoneme (see Appendix E for an example of each of the above two cases). Overall, and in agreement with PMSP96's claim, it appears that the factor that is at the root of the so-called regularity effect is not the existence of many words including a concurrent (regular) grapheme—phoneme correspondence as claimed in dual-route models (Coltheart & Rastle, 1994) but the existence of neighbors of high frequency that disagree with the target.

The position-of-irregularity effect on naming latency. Coltheart and Rastle (1994) argued that the position-of-irregularity effect they demonstrated for English polysyllabic words could be predicted only within the framework of dual-route models and constituted a challenge for current PDP connectionist models. This effect consists of a decrease in the regularity effect as the position of irregularity is farther from the beginning of the word. This effect was first pointed out by Jared and Seidenberg (1990), who obtained a much stronger regularity effect for the first syllable than for the last syllable of low-frequency bi- and tri-syllabic words. In a post hoc analysis of the regularity effects in a naming latency experiment in French, Content (1991) also found that the position of irregularity in a word interacted with the size of the regularity effect. However, this effect was not replicated in French (in another post hoc analysis) by Content and Peerenboom (1995), who reported a nonsignificant trend for regularity effects to be higher when the relevant phoneme appeared early in irregular words than when it occurred late; they further observed that regularization errors more frequently occurred when irregularity occurred at the beginning of the word than at the end. Unfortunately, it appears to be almost impossible

**Figure 12.** A simulation of the regularity by frequency by neighborhood consistency interaction in which high-frequency values have been artificially fixed (see text). IRR = irregular words; REG = regular words.

ported in the empirical studies conducted in English is found here only with respect to words with frequent enemies. This probably follows from the fact that most naming latency experiments in English (as in Taraban & McClelland, 1987) are restricted to monosyllabic words and that monosyllabic low-frequency exception words typically have enemies of high frequency in their neighborhood.

In this respect, it is noticeable that the latency pattern simulated on words with frequent enemies is comparable with the latency pattern of Taraban and McClelland's (1987) experiment (an experiment whose results are simulated by SM89 and PMSP96). In particular, it is found that, as in the experimental data, the mean naming latency of low-frequency regular words is higher than the mean naming latency of high-frequency irregular
to design a rigorous experiment specifically addressing this effect of irregularity position in French; given that irregular words are rare, it is not obvious how a strict and systematic manipulation of the position of irregularity could be done. A theoretical demonstration of this effect was attempted by “creating” virtual irregularities in otherwise regular words. For a given regular word from the database, the “regular” pronunciation of one grapheme was replaced by an irregular one during learning. The irregular pronunciation was generated by successively substituting, one at a time, each “regular phoneme” by another phoneme, thus creating an unexpected or “exceptional” correspondence for each phonemic position within the word. For example, for the printed word *radical*, whose appropriate (regular) phonology is /Radikal/, the “irregular” phonological codes /$adi$kal/, /$dikal$/, Ra$ikal$/, /Rad$kal$/, /Rad$al$/, and /Radik$s$/ (where /$$/ was put as a virtual unexpected phoneme) were created. Naming latencies were calculated for each irregular form and compared to the naming latency of the perfectly matched regular form. For each couple, the latency difference between the regular and the irregular form was taken as a measure of the position-of-irregularity effect.

Two sets of high-frequency and low-frequency words, learned by the network, were selected according to the following criteria: Words are all accurately read globally by the network, are tri-syllabic regular words of six to seven letters, and have a single consonant as onset and a syllabic structure of the type C1V1cC2V2cC3V3c (with “c” denoting here an optional consonant phoneme). Thus, computing the latency differences only for the C1, V1, C2, V2, C3, and V3 phonemes allows a strict comparison of the regularity effect between the three syllables. The experimental sets consisted of 76 low-frequency words (mean frequency = 156, range = 51–221) and 68 high-frequency words (mean frequency = 7,035, range = 1,922–48,163). For a given regular word, each virtual irregular form was learned by the network (thus replacing the regular form) and next presented in the test. When at least one irregular form yielded an erroneous response, the word was not taken into account for the subsequent latency analysis. For each syllable, the mean error rate and the mean naming latency difference between irregular and regular forms of each word was calculated collapsing the C and V phonemes (thus contrasting C1V1, C2V2, and C3V3). The results are given in Figure 13 for each high-frequency and low-frequency list and each syllable.

The main result shows that for low-frequency words, the regularity effect decreases very significantly as the position of irregularity moves farther from the first to the third syllable of the word, $F(2, 266) = 85.14, p < .001$. This gradient is not found for high-frequency words, $F(2, 266) < 1$, and the Frequency × Position interaction is highly significant, $F(2, 266) = 36.5, p < .0001$. The results also show the absence of a trade-off between naming latency differences and error rates. Rather, the error rate pattern is strictly similar to the latency pattern. In particular, regularization errors more frequently occurred when the position of irregularity was at the beginning of the word, in agreement with Content and Peereman’s (1993) experimental findings.

The gradient of regularity was therefore simulated in this virtual experiment as was the Regularity × Frequency interaction. Coltheart and Rastle (1994) argued that this gradient could only be explained within the framework of dual-route models as resulting, at least for low-frequency words, from a conflict at the phonemic stage between lexical and non-lexical phonemic information, . . . . If non-lexical phonemic information reaches that stage serially, left-to-right, it obviously follows that the later in the phonemic string the conflicting information lies, the higher the probability that lexical processing will be completed before the conflicting information reaches the phonemic stage. (p. 1204)

It is demonstrated here that the existence of a position-of-irregularity effect cannot be accounted for only by dual-route models. Within the present framework, these effects do not result from a conflict between two, lexical and sublexical, parallel procedures. On the contrary, they emerge within the global reading mode itself. From a theoretical viewpoint, it is noteworthy that a serial-like phenomenon (the regularity gradient) can be simulated here without resorting to any serial processing, against the claim of dual-route models. In the present model, the regularity gradient follows from the alignment principle according to which words match EM word traces by alignment of the first vocalic grapheme of the first syllable. As a direct consequence, when a regular word is presented to the network, those word traces that are significantly activated within the EM layer are more likely to be orthographically similar to the target on the first syllable than on any other syllable. Thus, an activated word trace has a higher probability for matching the first syllable of the target word than any other subsequent syllable. Given that all word traces activated within the EM layer contribute to phonological activation, many more word traces reinforce the activation of the relevant (regular) phoneme when belonging to the first syllable of the word than to other syllables. When an irregularity is virtually created on a given phoneme, the same words that reinforce the regular phoneme now become enemies for the irregular one. Thus, the initial phonemes that had more
that the number of episodes experienced with each word is the relative frequency of words remains the same over time but you become met at a given time, was manipulated. It was assumed that the relative frequency of words remains the same over time but that the number of episodes experienced with each word is globally reduced in less skilled (or younger) readers. Reading acquisition per se will not be simulated here; indeed, the following simulations were conducted using the overall database of 13,165 words, then minimally simulating the performance of a reader who has already encountered each of these words. A true simulation of beginning reading would have to start with a very reduced database gradually extended by the addition of new items; this has not been done here and will be the object of future developments of the network. For the present purpose, variations in reading skill were simulated by modifying the value of the \( \alpha_T \) parameter (fixed to \( 10^{-3} \) in the previous simulations of skilled reading).

Taking into account Equation 7, which defines parameter \( \alpha_T \), the number of occurrences, expressed in Equation 6, of a given word \( i \) at time \( T \) can be written as \( z_i(T) = \alpha_T f_i \). Thus, a reduction of parameter \( \alpha_T \) by any factor will correspond to an equal reduction of the number of word occurrences. Up to now, the network was run with \( \alpha_T = 10^{-3} \), thus simulating the performance of a skilled reader. An initial simulation was run with \( \alpha_T = 10^{-4} \), thus simulating the performance of somebody who experienced words 10 times less often than “our skilled reader.” We performed a second simulation with \( \alpha_T = 10^{-5} \), thus simulating a still less skilled reader who would have seen a hundred times fewer words than “our skilled reader.” We further simulated an “ideal” reader with a very high degree of expertise with \( \alpha_T \) tending towards infinity. A manipulation of the \( \alpha_T \) parameter value thus gives us the opportunity to estimate how reading ability develops over time with experience. Note that because the value taken by parameter \( \alpha_T \) determines the word frequency factor \( \varphi \), expressed in Equation 9, a decrease in \( \alpha_T \) value involves a decrease in frequency compression. For example, the already defined compression indicator \( \left( \varphi_{\text{max}}/\varphi_{\text{med}} \right) \) becomes equal to 1, 2, 2, 17.1, and 167.1 for \( \alpha_T = +\infty \), \( \alpha_T = 10^{-3} \), \( \alpha_T = 10^{-4} \), and \( \alpha_T = 10^{-5} \), respectively.

**Modulation of Reading Mode With Experience**

A simulation was done on a sample of 576 polysyllabic words randomly selected from the original database in order to determine how performance varies as a function of \( \alpha_T \) values. This sample, which is representative of the overall database, included words from one to five syllables and 1 to 10 letters whose frequency was between 8 and 2,010,176 occurrences per 100 million (\( Mdn = 501.5 \)).

The simulation shows (see Table 8) that the number of words processed globally increases with the value of \( \alpha_T \) and that the probability that a word will be processed globally remains an increasing function of its frequency (Figure 14). This suggests that an unskilled (or younger) reader should process many more words analytically, the probability for a word to be read globally increasing with experience and word familiarity. These simulations suggest that, when processing words, readers move on from the use of both analytic and global strategies to the preferential use of the global strategy. This is reminiscent of the hypothesis of models of reading acquisition (Chall, 1983; Frith, 1985) that learning is characterized by the establishment of an alphabetic stage followed by an orthographic stage. According to these models, the alphabetic and orthographic skills are acquired in strict sequential order but the alphabetic strategy remains available (although being less accessible) once the orthographic strategy has become established. The current simulation gives insights on the way the two strategies could differentially...
apply at relatively late learning steps. In particular, it is suggested that the two strategies should be simultaneously found in a given (not very skilled) reader when processing words of varying degrees of familiarity.

It is furthermore noteworthy that changing the value of $\alpha_T$ has an impact on the rate of global and analytic processing but leaves almost constant the number of words that are correctly processed in each reading mode. Indeed, 99.65% of words were correctly read at $\alpha_T = +\infty$, 98.44% at $\alpha_T = 10^{-3}$, 94.49% at $\alpha_T = 10^{-4}$, and 94.10% at $\alpha_T = 10^{-5}$. In turn, this result clearly demonstrates that the choice of $\alpha_T$ value is not crucial to the network’s capacity to read words with a high level of accuracy. This fact is worth noting in that the problem of the sensitivity of the parameter choice on simulation results is a central question in connectionist modeling.

Because analytic processing takes place only after global processing has failed and because the probability that a word will be processed analytically decreases with experience, the model predicts naming reaction times to decrease with experience. This prediction is consistent with data (Backman, Bruck, Hébert, & Seidenberg, 1984) showing that less experienced (younger) readers name words at longer latencies than (older) more skilled readers.

**Modulation of the Regularity Effect With Reading Experience**

Studies of the performance of skilled readers show that regularity effects are larger for lower frequency words and tend to disappear in higher frequency words (Andrews, 1982; Seidenberg et al., 1984; see also Content, 1991, for French words). Studies of the regularity by frequency interaction also demonstrate the existence of individual differences among readers. Seidenberg (1985) found that very skilled adult readers (fast-reader group) show no regularity effect even for low-frequency items, whereas less skilled adult readers (medium- and slowest reader groups) demonstrate a significant regularity by frequency interaction. The regularity effect therefore primarily reflects the word reading skill level of the participants. This is supported by developmental studies (Backman et al., 1984) that demonstrate that young readers show larger regularity effects on error rates than more skilled readers. These studies further establish that children in the early stages of reading acquisition perform poorly even on higher frequency exception words. The effect of reading experience is therefore to decrease the error score to a point at which regularity effects disappear on higher frequency words but remain on lower frequency items.

The network’s ability to account for the acquisition of naming skills has been assessed by analyzing its performance in naming regular and irregular words at different word experience levels (i.e., different $\alpha_T$ values). The simulation was conducted for each $\alpha_T$ value on a sample of 566 regular and 114 irregular mono- and polysyllabic words. Irregular words were divided into two groups of 57 lower frequency words (median frequency = 131, range = 4-544, mean frequency = 183, SD = 155) and 57 higher frequency words (median frequency = 2,641, range = 604-85,992, mean frequency = 7,478, SD = 13,306). The regular words were also divided into two subsets with the same frequency cut-off (574), a first set of 295 regular words of lower frequency (median frequency = 221, range = 8-570, mean frequency = 254, SD = 126) and a second set of 271 regular words of higher frequency (median frequency = 2,127, range = 587-2,010,176, mean frequency = 17,333, SD = 131,260). The results are presented in Figure 15.

The network’s performance at different $\alpha_T$ values parallels human behavioral data. At $\alpha_T = 10^{-4}$, a regularity effect is obtained for both high-frequency and low-frequency words, regular words producing a smaller error rate than irregular words. The regularity effect is progressively reduced for high-frequency words from $\alpha_T = 10^{-4}$ to $\alpha_T = 10^{-3}$ and disappears completely when $\alpha_T$ tends towards infinity. In contrast, a regularity effect does emerge at all $\alpha_T$ values for low-frequency words, although being progressively reduced from $\alpha_T = 10^{-3}$ to $\alpha_T$ tending towards infinity. It is also noteworthy that more than 70% of the errors made on low-frequency exception words are regularizations, whatever the $\alpha_T$ values. Overall, the model satisfactorily replicates the evolution of the regularity by frequency interaction as a function of reading experience as previously shown in the SM89 model.

**Acquired Dyslexia**

Any plausible reading model must offer a reasonable explanation of how the different forms of acquired dyslexia arise and particularly of the double dissociation observed between phonological and surface dyslexia. Phonological dyslexia is characterized by the selective impairment of pseudoword reading and relative preservation of regular and irregular word reading (Beauvois & Derouesné, 1979; Patterson, 1982; Shallice & Warrington, 1980). In contrast, irregular word reading is selectively impaired in surface dyslexia relative to the reading aloud of regular words and pseudowords (Marshall & Newcombe, 1973; Patterson et al., 1985). This double dissociation has received different interpretations within the framework of previous reading models. Some models that postulate the existence of a semantic and a "phonological" pathway (multiple-level model: Shallice, 1988; Shallice et al., 1983; and the SM89 general reading model) emphasize the role of the semantic route in the explanation of these acquired dyslexias. A preservation of this pathway alongside with an impairment of the phonological route (or a deficit of phonology per se) is hypothesized to account for phonological dyslexia, whereas surface dyslexia is seen as resulting from damage to the semantic pathway and only partial damage to the phonological route. The interpretation of these two forms of acquired reading impairment is somewhat different within the framework of triple-route models. In the context of these latter models, phonological dyslexia is viewed as resulting from an impairment to the nonlexical pathway with a preservation of at least the lexical nonsemantic pathway, whereas surface dyslexia might reflect a disorder of the lexical pathways with the preservation of the nonlexical route. It is therefore assumed in these models that the performance of individuals with phonological dyslexia does not necessarily follow from processing through the sole semantic pathway. As a consequence, these two classes of models make different predictions about normal and pathological reading (see Carbonnel & Ans, 1996, for a review). The first class of model predicts that the pattern of phonological dyslexia cannot be observed in association with a severe impair-
ment of the semantic pathway. In contrast, triple-route models predict that phonological dyslexia and surface dyslexia can both occur in association with a severe impairment of the semantic pathway. Our aim in this section is to provide an account of the double dissociation characteristic of phonological and surface dyslexia within the implemented orthography-to-phonology network. It is shown that these two patterns of dyslexia can be obtained following distinct functional lesions of the network.

**Phonological Dyslexia**

The challenge here is to demonstrate that pseudoword reading can be selectively impaired following specific damage to the implemented network. Furthermore, a strong dissociation between word and pseudoword reading must be demonstrated because many cases of phonological dyslexia have been reported (e.g., Beauvois & Derouesné, 1979; Carbonnel, David, Charnallet, & Pellat, 1987; Coslett, 1991; Funnell, 1983; Sartori, Masetterson, & Job, 1987) in which an almost perfect performance in word reading dramatically contrasts with very poor performance with pseudowords (0% to 20% of correct pronunciations). In the present network, two kinds of impairment can result in the pattern of performance characteristic of phonological dyslexia.

Phonological dyslexia could first originate from an impairment to the orthographic checking process. In such a case, no mismatch between O1 and O2 could ever be detected, so that any O2 orthographic echo would always be considered as correct. Consequently, whatever the input stimulus, the reading system will never shift to the analytic mode. The phonological output will then always result from global processing. Thus, the result of disrupting the orthographic checking process could be simulated by running the network in a forced global mode. Using this technique, we assessed reading performance on a list of 566 regular words (from the previous section), a list of 92 irregular words matched in median frequency and frequency range to the regular words, and the list of 830 pseudowords.

---

**Figure 15.** A summary of the regularity effects according to $\alpha_r$ values. IRR = irregular words; REG = regular words.
The results showed that 95.76% of regular words and 93.48% of irregular words were read accurately against only 27.71% of pseudowords. A detailed analysis of the model’s performance on pseudowords revealed that lexicalization errors frequently occurred. These errors were more often observed on those pseudowords that were very close (by one letter) to real words (390 of 655, or 59.45% erroneous responses, from which 86.09% were lexicalizations) than on more dissimilar (but legal) pseudowords (46 of 174, or 26.44% errors, among which 31.29% were lexicalizations). It therefore appears that any damage that would prevent shifting to the analytic mode would result in a pattern of reading performance similar to that typically reported in severe forms of phonological dyslexia. It is here hypothesized that poor pseudoword reading is caused by an impairment to an orthographic processing component (the orthographic checking process) of the reading system, thus preserving phonological processing. Such an account is compatible with the case of a French individual with phonological dyslexia, LB (Derouesné & Beauvois, 1985), who showed very poor performance in pseudoword reading (30% correct) but relatively preserved word reading (from 74% to 98% correct) in the context of very good phonological abilities (see Coltheart, 1996, for a discussion).

We therefore attempted to simulate the pattern of performance of patient LB on four of the pseudoword lists on which he was originally tested.16

As Figure 16 shows, the network's pattern of performance paralleled that of case LB. The network's performance was globally inferior to LB's but showed the same dissociation between visually similar and dissimilar pseudowords. The model therefore offers a plausible account of the pattern of phonological dyslexia without associated phonological disorders.

However, cases of phonological dyslexia without impairment of phonological processing seem to be the exception. In fact, all other reported cases of phonological dyslexia (e.g., Patterson & Marcel, 1992; Patterson, Suzuki, & Wydell, 1996; Farah, Stowe, & Levinson, 1996; Berndt, Haendiges, Mitchum, & Wayland, 1996) rather show a phonological deficit that is viewed as responsible for both poor nonword reading and poor performance on purely phonological tasks of phonemic awareness.

In the current general model, it is argued that analytic processing involves the sequential production of phonological segments (typically syllables) that are successively stored in a phonological buffer (not implemented), where they are temporarily maintained for subsequent blending. It would therefore be expected that a deficit of the blending process at the phonological buffer level would result in a selective difficulty in pronouncing those letter strings that are processed analytically (Hulme, 1987; Hulme & Mackenzie, 1992, for a similar suggestion in developmental dyslexia). It follows that pseudoword reading would be selectively affected. Thus, the blending deficit here hypothesized is viewed itself as resulting from an impairment to a component of the phonological processing system (i.e., the phonological working memory). This component is not specifically dedicated to reading but is also involved in some purely phonological tasks that require phonemic information to be temporarily maintained for processing. Thus, a deficit at this level should also lead to poor performance in metaphonological tasks as previously observed in most cases of phonological dyslexia. It is assumed in the current model that this phonological component plays a crucial role in reading (and in particular in reading acquisition without feedback) and its disorder offers a plausible explanation of the classic pattern of phonological dyslexia (poor nonword reading with impaired phonological processing and a short-term memory deficit).

Surface Dyslexia

The aim of this section is to demonstrate that damage to the network could also result in the selective impairment of irregular word reading as observed in surface dyslexia. An accurate simulation of this reading disorder would further require regularization errors to predominate because individuals with surface dyslexia typically mispronounce irregular words by giving a more regular pronunciation (e.g., *outil* /uti/ → /util/ or *orchidée* /oRkidé/ → /oRsidé/).

The present model should also account for the existence of different subtypes of acquired surface dyslexia. Indeed, a dissociation between regular and irregular word reading has been reported in some patients in association with a slowing of reading speed (Kay & Patterson, 1985; Shallice & Warrington, 1980). Apart from regularizations, other errors occur that sometimes result from the misapplication of GPC rules or from an almost total failure to apply the contextual rules (Henderson, 1982; Marcel, 1980). Also, purely visual errors are found to occur in some patients (Coltheart et al., 1983). This first subtype of surface dyslexia, referred to as nonfluent surface dyslexia (or word-form dyslexia), has been interpreted as arising from a relatively peripheral disorder (at the level of the early visual analysis of the orthographic input or at the word-form system level), an interpretation close to that proposed for letter-by-letter reading (see Shallice, 1988, for discussion). In a second

---

16 Only the lists of nonhomophonic pseudowords were retained because the pseudohomophone effect found in this patient was probably due to the influence of top-down processes not implemented here.
subtype of surface dyslexia, sometimes referred to as fluent surface dyslexia or semantic dyslexia, patients read regular words better than irregular in the context of fluent reading (Bub, Cancelliere, & Kertesz, 1985; McCarthy & Warrington, 1986). This form has typically been reported in patients with a severe semantic deficit for whom a disruption of the semantic pathway could be hypothesized. A regularity by frequency interaction has also been noted, participants being more prone to make errors on low-frequency irregular words than on high-frequency irregular words. This second subtype is viewed as resulting from a central impairment involving the lexical pathways.

With respect to the present model, acquired surface dyslexia could result from damage at two different levels. Nonfluent surface dyslexia could arise from a reduction of the visuo-attentional window through which information from the orthographic input is extracted. It was expected that within the present network such a reduction would lead the O2 echo to frequently be different from the O1 input. It has been verified that this was indeed the case, whatever the size of the portion of a word presented over O1, so that the system typically shifted to the analytic mode. It was also expected that the model’s performance would vary as a function of window size when running in analytic mode. Some reduction of the window size would preserve processing of syllable units within their context, thus activating EM segment traces in a normal way. More severe visuo-attentional damage would prevent the network’s taking into account the preceding (and already processed) orthographic context of the remaining segment to be read. Simulations of these two kinds of deficit were conducted in forced analytic mode with a normal (cs = 1) and restricted (cs = 0) context size parameter, respectively. This was done on the sample of 114 irregular words (from the Developmental Perspective section) on the list of 566 regular words and on a list of 100 pseudowords randomly selected from the 830 original items. The results are presented in Table 9. They show that errors typically concern irregular word processing when the system runs in forced analytic mode.

Regularizations were the dominant error type because they represented 92.86% (39 of 42) and 91.80% (56 of 61) of misreadings with cs = 1 and cs = 0, respectively. The error rate increased when the context size was decreased, and more items were read at the graphemic level with cs = 0 (84 of 780, or 10.77%) than with cs = 1 (44 of 780, or 5.64%). Furthermore, apart from regularization errors, 39.68% (25 of 63) of the errors could be viewed as resulting from the misapplication of contextual rules. They were all observed in regular word reading at cs = 0, whereas none occurred at cs = 1. These errors followed from the misapplication of the S rule; in French, s is pronounced /zb/ between two vowels, /s/ otherwise. When the context was restricted to zero, errors such as ‘adverse’ /advBr’s/ → /advBr’s/, ‘grisier’ /gRis’/ → /gRis’/, ‘prèsevier’ /pRèzéRvé/ → /pRèzéRvé/, ‘stire’ /siR’/ → /ziR’/ or ‘pension’ /pázjé/ → /pázjé/ were observed. Therefore, when running in forced analytic mode, the performance of the network showed most of the defining features reported in patients with a nonfluent surface dyslexia. The network’s performance was further compared with that of GR (Valdois, Tainturier, David, & Pellat, 1996), a French individual with nonfluent surface dyslexia exhibiting an associated visual attentional disorder. This participant in particular showed a strong regularity effect when words and pseudowords were presented in mixed lists. In this condition, 95% of the regular words (19 of 20) were read accurately against only 35% of the irregular words (7 of 20, matched in length and frequency). The simulation conducted on the same items with cs = 0 yielded the same global results (95% correct against 35% for the regular and irregular words, respectively). In both cases, all erroneous responses were regularizations. The performance of the network when running in forced analytic mode with a restricted context size parameter therefore seems consistent with the performance of this participant with nonfluent dyslexia.

Finally, our interpretation is compatible with the hypothesis of a continuum from surface dyslexia to letter-by-letter reading (Shallice, 1988). An extremely severe deficit of the visuo-attentional system could indeed prevent processing at the syllabic level. A simulation of this deficit would then grossly correspond to reading at the graphemic level.17 This simulation has not been conducted but could clearly lead to extreme regularizations such as houx /u/ pronounced /uks/ or poil /pjól/ pronounced /poil/ as sometimes observed in letter-by-letter readers. On this view, nonfluent surface dyslexia and letter-by-letter reading could both originate from the same kind of impairment, namely, a reduction in the size of the visuo-attentional window.

However, cases of fluent surface dyslexia do not exhibit a peripheral disorder of the kind previously hypothesized (Bub et al., 1985; McCarthy & Warrington, 1986). Their pattern of performance could be predicted within the current framework as resulting from more central damage to the connections linking the EM units to O2 units. Damage to these EM—O2 connections should lead the O2 input activation to be systematically lowered.18 For some words, in particular for some low-frequency

---

Table 9

<table>
<thead>
<tr>
<th>Pseudowords</th>
<th>Regular words</th>
<th>Irregular words</th>
</tr>
</thead>
<tbody>
<tr>
<td>cs</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>1</td>
<td>3/100</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>6/100</td>
<td>6</td>
</tr>
</tbody>
</table>

---

17 Behrmann and Shallice (1995) reported data that are in apparent contradiction to the present explanation of letter-by-letter reading. Indeed, they interpreted the deficit of their patient DS as resulting not from a visual spatial disorder but rather from a deficit of individual letter activation. As a consequence of this disorder, DS can only identify one letter at a time. With respect to the present model, such a deficit would have similar consequences to a severe reduction of the visuo-attentional window, with reading relying on graphemic processing.

18 Although a single macro EM word unit is assigned to a given word in the present implementation, this is only a formal equivalence that in fact represents a set of elementary word traces spread over layer EM. Consequently, in global mode, a partial disruption of EM—O2 connectivity does not completely suppress the contribution of corresponding words within the orthographic echo but only reduces the different contributions (in O2 activation) of all learned words.
words, read globally by the undamaged system, O2 unit activations would no longer reach the threshold (set to $\theta_0 = 0.35$ in this mode). In this case, the O1–O2 matching check would not succeed and the system would shift to the analytic reading mode. Regularization errors are then expected on these words when irregular but in the absence of contextual errors (as previously demonstrated when $cs = 1$). An analytic reading shift is all the less likely to occur than word frequency increases because, even with a lowered O2 input activation, the probability that the O2 threshold will be exceeded increases with word frequency. With respect to regular words, shifting in the analytic mode should have little effect, if any, because it has been demonstrated that regular words were correctly read in this mode. A regularity by frequency interaction is therefore expected, as reported in Bub et al.’s (1985) typical case. In order to demonstrate the potential of the current model in accounting for the pattern of fluent surface dyslexia, two simulations were conducted with a lesioned network. The network was damaged by removing each EM–O2 elementary connection with a probability of either $p = .4$ or $p = .6$. To evaluate the effect of lesions, we tested the network on the lists of regular words ($N = 566$), irregular words ($N = 114$), and pseudowords ($N = 100$) previously used to simulate nonfluent surface dyslexia. Several frequency ranges were required because it was predicted that the lesion effect should gradually decrease when irregular word frequency increased. It was not possible to find among the original lists (in particular of irregular words) a sufficient and equivalent number of items belonging to contrasted enough frequency ranges. Consequently, for each word category, the frequency of the whole set of words was arbitrarily fixed so that the same words were tested at each frequency range. Four frequency ranges were considered: low ($f = 200$ per 100 million), medium-low ($f = 500$), medium-high ($f = 1,000$), and high ($f = 5,000$). Simulations were thus conducted with the intact ($p = 0$) or lesioned ($p = .4$ and $p = .6$) network on the same lists of regular and irregular words assigned with a virtual frequency of 200, 500, 1,000, or 5,000. Results showed that a regularity by frequency interaction was obtained from the two lesioned networks. As expected, the lesions have virtually no effect on regular word reading (Figure 17, left part). In contrast, for irregular words, the lesion (especially for $p = .6$) induces an impressive and gradual increase of the error rate when frequency decreases (Figure 17, middle part). The same pattern was observed when regularization errors alone were considered (Figure 17, right part) because most errors (from 90 to 100% at $p = .6$) were regularizations, whatever the frequency range, as also expected. Finally, the lesion (at either $p = .4$ or $p = .6$) did not affect pseudoword reading as typically found in fluent surface dyslexia.

Overall, the main subtypes of phonological and surface acquired dyslexia can be accounted for following specific damage to the network. In particular, the model predicts that phonological dyslexia can result either from damage to a purely orthographic component or from damage to a purely phonological component. This account therefore offers a plausible explanation of the case of phonological dyslexia without associated phonological impairment as well as those cases that show a phonological disorder. The model further predicts that a visual input disorder could be at the origin of some forms of surface dyslexia, whereas a more central deficit could account for other surface dyslexic patterns. This gives new insights into the distinction sometimes made in the literature between nonfluent and fluent surface dyslexias.

General Discussion

The current connectionist network adopted the fundamental proposal of Hintzman (1984, 1986, 1988) that during learning, experienced events give rise to localized traces whereas the whole set of traces contributes to retrieval. After training on an extensive database composed of both mono- and polysyllabic French printed words and their syllabic segments, the network was able to generate the correct pronunciation of mono- and polysyllabic regular words, irregular words, and pseudowords.

![Figure 17](https://example.com/figure17.png)

*Figure 17.* Simulation results accounting for the error pattern of fluent surface dyslexia. Two lesioned networks ($p = .4$ and $p = .6$) and the intact network ($p = 0$) are considered, where $p$ denotes the probability that an elementary EM–O2 connection will be removed. Left and middle: Error rate on regular and irregular words, respectively. Right: Rate of regularization errors (with respect to the number of items).
at a level of accuracy comparable to that of skilled readers. The network also simulated the main effects of frequency and consistency on naming latency as well as the frequency by consistency interaction. The effect of reading experience on naming accuracy was also investigated; early in training, the network performed more poorly on both low- and high-frequency irregular words as observed in less skilled readers. Finally, the patterns of acquired phonological and surface dyslexia were simulated after specific damage to the network.

The model implemented in the current work develops a view of the reading system that differs from previous ones in a number of respects. Following SM89 and PMSP96, the model does not retain the idea incorporated in the dual-route model (Coltheart et al., 1993; Reggia et al., 1988) that a sublexical procedure involving the use of a conversion rule system is required for pseudoword reading. An alternative view was proposed by SM89 and PMSP96, who demonstrated that a network that learns gradually by capturing in its connection weights the statistical structure among orthographic and phonological correspondences was able to read pseudowords as well as regular and irregular words. Nevertheless, the current model also differs from this alternative view: Learning was localized so that spelling–sound regularities are not captured in the connection weights during learning. The pronunciation of an orthographic input emerges only at the time of retrieval from the integrated activation of localized traces. As a consequence, our interpretation of the consistency effects differs from that of both dual-route and PDP models in that those effects exclusively result from conflicts between an always limited number of localized traces (corresponding to the target word and its closest neighbors) and not from conflicts between outputs from separate routes or interference between knowledge of spelling–sound correspondences implicitly captured in the connection weights. It straightforwardly follows from these underlying principles that, at least for low-frequency words, naming latency is entirely governed by neighborhood consistency (in particular, presence or absence of enemies) as shown in the simulation of Peerman’s (1995) experiments and in line with Jared and Seidenberg’s (1990) data. In the same way, the network provides an account of the position-of-irregularity effect on naming latency (Coltheart & Rastle, 1994). As first hypothesized by PMSP96, this effect was interpreted here as a consistency effect that derives from the fact that word traces that are activated by the printed input and agree or conflict with its pronunciation are more likely to be orthographically similar to the target on the first syllable than on any other syllable.

Another feature of the current model is that the pronunciation, in global reading mode, of any novel word can be generated only from the integration of the orthographically similar word traces activated by the input (reverberation process). As a consequence, the generalization power of the current network is more limited than that of the connectionist networks of PMSP96. In particular, the ability of the network to generalize to polysyllabic pseudowords is limited because these items are statistically less likely than monosyllabic stimuli to have orthographic neighbors liable to combine for generating their correct pronunciation. The limited power of generalization from word trace activation is viewed as a strength of the present network because experimental findings (Carbonnel et al., 1998) suggested that pseudowords, and in particular polysyllabic pseudowords, involve a reading procedure that typically differs from that which supports familiar word reading. It is here assumed that an analytic reading procedure that rests on the activation of segment traces rather than word traces typically supports pseudoword reading and in particular polysyllabic pseudoword reading.

The two global and analytic reading procedures implemented here do not work in parallel as hypothesized in the dual-route model. The global procedure always proceeds first, the analytic procedure being used only after global processing has failed. A strong prediction of the assumption that the sequential (analytic) procedure follows global processing and supports the reading of a great number of polysyllabic pseudowords is that the naming latency of all familiar words would be systematically shorter than the naming latency of any unfamiliar word. In agreement with this prediction, Carbonnel et al. (1998) found that the latency distribution of words and pseudowords was largely bimodal with a slight overlap for monosyllabic stimuli but no overlap at all for bi- and tri-syllabic words and pseudowords.

Other Empirical Issues

Coltheart et al. (1993) claimed that any serious model of reading aloud should answer six important questions. In particular, the model should give a reasonable explanation of irregular word and pseudoword reading, of lexical-decision results and of acquired and developmental dyslexia. They concluded that the dual-route model alone offered a satisfactory answer to all six questions, claiming that no single route model of reading would successfully answer them all. PMSP96 challenged this view by attempting to demonstrate that their model reasonably addressed these issues. The ability of the present model to do so is now discussed. Issues concerning regular word, irregular word, and nonword reading have been addressed extensively in the current work. We now consider in more depth how neuropsychological findings support our account of acquired phonological and surface dyslexia, how the model succeeds in providing an account of the different subtypes of developmental dyslexia, and its predictions with respect to lexical-decision performance. It should be emphasized that the current model alone successfully answers a seventh and crucial question: How do skilled readers read polysyllabic words and pseudowords aloud?

Acquired Phonological and Surface Dyslexias

In accordance with the dual-route (in fact triple-route) and multiple-level models, it is argued here that any kind of isolated word (regular word, irregular word, and nonword) can be accurately read without resorting to semantics. Indeed, in the present model, all words can be reliably translated from print to sound through a fully competent phonological pathway as already hypothesized in the multiple-level model. The current model differs from the triple-route model in postulating a single pathway able to convert any letter string into sound, whereas this latter model supposes, apart from the semantic route, the existence of two additional pathways: a nonlexical pathway and a lexical assemblage pathway dedicated, respectively, to nonword processing and irregular word reading. Finally, although postulating a single phonological pathway, the PMSP96 model strongly differs from
the present one (and all other cognitive models of reading) in postulating that semantic support is normally required for naming (at least) some irregular words (see also Hillis & Caramazza, 1995). These different conceptions of the normal reading system lead to different predictions about the origin of acquired dyslexia. In the present model, the patterns of phonological and surface dyslexia can both emerge from impairments within the phonological pathway itself, so that either of these two forms of acquired dyslexias is expected to sometimes occur in association with a severe semantic deficit. This latter prediction is consistent only with the triple-route model and is in sharp contrast to predictions from both the multiple-level model and the PMSP96 model. More schematically, three types of model (the present one, the triple-route model, and the multiple-level model) predict the possible preservation of all-kinds-of-word reading following a severe impairment of the sole semantic pathway, whereas the fourth simulation of the PMSP96 model (see also Patterson & Hodges, 1992) predicts that an impairment to semantic representations should be accompanied by surface dyslexia. Two types of model (the present one and the triple-route model) further predict that either surface dyslexia or phonological dyslexia could follow from a severe impairment of the semantic pathway associated with some impairment of the nonsemantic pathway. In contrast, the multiple-level model and the PMSP96 model (see also Hillis & Caramazza, 1995) predict that surface dyslexia but not phonological dyslexia will follow from such an impairment. As a consequence, the pattern of phonological dyslexia is interpreted as resulting from a severe disruption of the phonological pathway and reading via semantics (multiple-level model and PMSP96 model). This disorder could also be interpreted within PMSP96’s model as following from an impairment to a more general phonological processing system that is involved in all phonological tasks and is thus not specific to reading (Farah et al., 1996; Patterson et al., 1996).

The current model, as well as triple-route models, does not need to postulate that the semantic pathway can accurately read all kinds of isolated real words (see Valdois, Carbonnel, David, Roussel, & Pellat, 1995, for a similar proposal in a case study of deep dysphasia) to account for phonological dyslexia (at least for the form in which the reading disorder is not associated with a general phonological impairment as in case LB; Dehouesné & Beauvois, 1985).

Three questions are therefore relevant to assess the predictions of these different reading models: Is a semantic disorder systematically associated with surface dyslexia? Does phonological dyslexia never occur in association with a severe semantic deficit? Is there empirical evidence that all words can be accurately read via the sole semantic pathway?

Empirical evidence concerning these three questions is now considered in order to determine which of these different conceptions of the normal reading system accounts for the broadest range of pathological data. The hypothesis that an impairment of a phonological processing system could be responsible for some forms of phonological dyslexia is also discussed.

Is a semantic disorder systematically associated with surface dyslexia? On the one hand, some cases, in particular WLP (Schwartz, Saffran, & Marin, 1980) and DRN (Cipolotti & Warrington, 1995), have been reported in which a severe semantic disorder occurs together with intact irregular word reading. As PMSP96 themselves acknowledged, these patients show the very pattern of impairment that their model excludes, namely, a semantic impairment without surface dyslexia. In order to accommodate these data, Plaut et al. (1996) attempted a post hoc interpretation of these cases in terms of individual differences, arguing that, in these individuals, the “phonological pathway had developed a relatively high degree of competence without assistance from semantics” (PMSP96, p. 99). In the absence of empirical data supporting this interpretation (and considering that such cases do not seem to be so exceptional), these cases will be taken here as support for the theoretical assumptions of the current model.

On the other hand, some other empirical data show the existence of a correlation between semantic impairments and surface dyslexia (Patterson, Graham, & Hodges, 1994); these data have to be carefully considered because they have been interpreted by PMSP96 as support for their model and could be viewed as challenging the current account. In their analysis of the performance of patients with some form of progressive dementia—dementia of the Alzheimer type (DAT), progressive fluent aphasia, or semantic dementia—Patterson and Hodges (1992), Graham, Hodges, and Patterson (1994), and Patterson et al. reported that poor performance on tasks designed for assessing semantic abilities was accompanied by poor performance in irregular word reading. Such a correlation was taken as empirical evidence in support of the PMSP96 reading model, because it could be viewed as showing that a semantic disorder caused surface dyslexia. In fact, the interpretation of Patterson and Hodges’s results is not really straightforward.

Indeed, as argued by Patterson and Hodges (1992) themselves, the association between poor comprehension and poor irregular word reading exists “not because the actual translation from orthography to phonology requires participation from word meaning; it is rather because the integrity of lexical representations depends on their normal interaction with the semantic system” (p. 1035). In other words, Patterson and Hodges (see also Graham et al., 1994; Patterson et al., 1994) claimed that a semantic deficit causes a degradation of the phonological whole word representations, which in turn results in the irregular word reading disorder—for example, according to Patterson et al., “The tendency to misread exception words follows as a direct consequence of this vulnerability of whole-word representations” (p. 405). It is also argued that such a degradation of the phonological output representations has no impact on the phonological processing abilities of these participants, who have normal performance in tasks of repetition and phonological awareness (Graham, Patterson, & Hodges, 1995). However, this interpretation requires a number of comments. First, in a very general way, it is clear that although a causal relation implies a correlation, a correlation does not imply a causal relation. As a consequence, the existence of a correlation between a semantic deficit and a degradation of lexical phonological representations (Patterson et al., 1994) gives no insights into the potentially causal nature of their link. Second, one could speculate that the results reported in Patterson and coworkers’ studies could reflect the reverse causal relation as well; the degradation of whole word representations being then viewed as responsible for the semantic deficit. In this respect, it is noteworthy that the patients described by Patterson et al. showed a performance that became
worse as the task placed stronger demands on output phonology (picture naming, naming to description, category fluency). In fact, the DAT participants were only slightly impaired in the sole task that required no speech output production (the word–picture matching test with close semantic foils) and that should have been expected to be the most sensitive to a semantic processing impairment itself. It is therefore questionable whether the results of Graham et al. (1994) and Patterson et al. constitute evidence that a semantic impairment invariably causes surface dyslexia.

Furthermore, counterproofs for such an implication do exist through cases that demonstrate a severe semantic disorder without associated surface dyslexia. WLP and DRN are such cases, but still more relevant are cases DC (Lambon-Ralph, Ellis, & Franklin, 1995) and MB (Raymer & Berndt, 1994). The performance of DC was explicitly investigated to assess the hypothesis of a contribution of semantics to normal reading. DC is a DAT patient who showed only occasional word-finding difficulties and relatively preserved spontaneous speech but exhibited a semantic disorder of a type and severity comparable to the patients studied by Patterson and colleagues (Graham et al., 1994, 1995; Patterson et al., 1994; Patterson & Hodges, 1992). An extensive investigation of DC’s reading performance showed essentially preserved abilities in irregular word reading (237 of 244 correct)—on tasks including even very irregular or abstract words—in spite of an inability to provide any accurate semantic information about these words. This case therefore shows that a severe semantic disorder is not systematically associated with selective difficulties in irregular word reading. In the same way, patient MB (Raymer & Berndt, 1994) demonstrated a preserved irregular word reading (100% correct for high- and medium-frequency exception words, 93% correct for low-frequency exceptions) despite chance-level performance on most semantic tasks. In fact, this patient seemed even more impaired on semantic tasks than the more severely impaired DAT patients reported by Patterson et al. (1994). Indeed, he succeeded on only 15% of items in oral naming, 2.3% in naming to definition, and 52% in word–picture matching (with semantic foils) against 53.5%, 25%, and 81.7% on similar tasks (although with different stimuli) for the patients reported by Patterson et al. The existence of these two last cases of patients with a severe semantic deficit without associated surface dyslexia is inconsistent with the hypothesis that a semantic disorder implicates surface dyslexia. Overall, the available neuropsychological data does not seem to provide clear evidence against the current account of surface dyslexia.

**Does phonological dyslexia never occur in association with a severe semantic deficit?** Coltheart et al. (1993) claimed that such an association does sometimes occur, on the basis of data from two cases, WB (Funnell, 1983) and WT (Coslett, 1991), who exhibited a pattern of phonological dyslexia associated with a severe semantic disorder. WB showed a sharp contrast between an almost preserved word reading (from 80% to 93% correct) and an almost complete inability to read pseudowords (1 of 30 correct). However, WB’s good word reading performance did not appear to be semantically mediated because his comprehension of words was poor on semantic matching tasks. WT (Coslett, 1991), on his own, exhibited a selective impairment in the comprehension of abstract words that he could nevertheless read accurately (95% correct). According to Coslett, intact abstract word reading (without comprehension) could not be viewed as relying solely on the nonlexical pathway because WT was able to read irregular words while being very poor on nonwords (20% correct). These results at first glance suggest a major deficit of the phonological pathway (responsible for poor pseudoword reading) together with a major impairment of the semantic pathway so that the preserved ability of these patients to accurately read irregular words would not be predicted by PMS96’s model. However, an attempt could be made to reconcile these cases of phonological dyslexia with PM96’s model by arguing that the documented word–nonword dissociation does not reflect a disorder of the phonological pathway per se but rather results from an impairment of a purely phonological processing system that is not specifically dedicated to reading tasks and is not part of the general model outlined in SM89. Thus, the phonological pathway could be sufficiently intact to support (with the help of even a partially damaged semantic pathway) good irregular word reading. Such an explanation is reminiscent of that advanced by Hillis and Caramazza (1991, 1995; see also PMS96) that the preserved word reading of phonological dyslexia would result from the summation of partial information from the impaired semantic and phonological pathways. Such a hypothesis implies that at least some partial information must be available from one or the other pathway, poor word reading abilities being expected when one of these two pathways is totally inoperant.

In this respect, Hillis and Caramazza (1995) reported cases of patients whose performance seems to support this prediction. However, other cases (Raymer & Berndt, 1994; Sartori et al., 1997) that provide evidence against the summation hypothesis (see Carbonnel, 1996) do exist. One of these cases is that of an Italian patient, Lisa (Sartori et al., 1987), who showed a severe semantic deficit and a phonological dyslexia (80% vs. 10% correct on words and nonwords, respectively) together with some language disorder (fluent speech without content but perfect repetition of words and sentences). The good word reading performance of Lisa cannot be interpreted as resulting from the summation of partial semantic and phonological information because her semantic knowledge about words she read correctly was most of the time totally abolished. For instance, she scored at chance level in comprehension tasks as simple as word–picture matching and an odd word out test (in which she had to indicate a single word that did not belong to the same semantic category as other items). In a task that involved only loose semantic knowledge, in which she had to sort 40 nouns into animal and nonanimal categories, her performance was also at chance level, although most words could be read accurately (34 of 40). Still more interesting, among the words accurately read, 18 were correctly categorized but 16 were not. In sum, Lisa’s comprehension was very poor and her performance in word reading was clearly independent from her comprehension.
of the same words. Overall, her good word reading can in no way be interpreted as relying on even partial semantic knowledge about words. Rather, there is clear evidence that her reading entirely relied on the phonological pathway, although it was characterized by a strong dissociation between performance with words and nonwords.

The current model straightforwardly accounts for Lisa's reading pattern as reflecting a selective preservation, within the phonological pathway, of reading in the global mode. Although Lisa's performance on both words and nonwords is slightly inferior to that of the current network when clamped in the global mode, the two patterns are similar. Particularly in agreement with our account is the observation that Lisa read correctly twice as many pseudowords that were visually similar to words as pseudowords that were not visually similar to words (16.2% vs. 6.9%) as was also observed in the network (27.6% vs. 13%). Moreover, most reading errors on visually similar pseudowords were lexicalization errors in the patient (58%) as well as in the simulation (59.4%). Finally, in lexical decision, she showed (when considering results summed across all lists) a higher rate of false alarms on visually similar pseudowords (60%) than on dissimilar pseudowords (30%), a rate very close to that of lexicalization errors in the simulation (59.4% vs. 26%). It must further be stressed that the performance of Lisa on pseudowords was as poor in visual lexical decision (in which no verbal output is required) as in reading. This finding, added to the fact that her repetition was totally preserved, argues against an explanation of the nonword reading deficit in terms of a purely phonological disorder. 

Another study of phonological dyslexia reported by Raymer and Berndt (1994) provides data that are even more in conflict with the summation hypothesis; these authors in fact explicitly designed their study with the aim to assess Hillis and Caramazza's (1991) hypothesis. Two patients, MB and SC, were examined on tasks of word comprehension, reading and spelling of the same words, and pseudoword reading. The two patients were very impaired on the comprehension tasks while showing a sharper contrast between word (96% and 68% correct for MB and SC, respectively) and pseudoword reading (67% correct for MB and 9% correct for SC). Neither patient showed any difference between regular and irregular word reading. The observation in these patients of a very severe semantic deficit (with chance-level performance on even very simple tasks such as picture–picture matching) associated, at least for MB, with intact irregular word reading is the first clear evidence against the summation hypothesis. But other aspects of the performance of these patients are still more problematic for this hypothesis. Indeed, Raymer and Berndt contrasted the patients' reading performance on two categories of irregular words: "known words," for which they gave evidence for at least partial semantic knowledge and "unknown words," for which performance was below chance level on semantic tasks. According to the summation hypothesis, reading of known words might be expected to be relatively preserved and reading of unknown words very poor. The results were inconsistent with this prediction because the patient's reading performance was comparable for the two word categories (97% and 94% correct for MB; 68% and 71% correct for SC). Therefore, these results clearly indicated no relationship between semantic knowledge and irregular word reading. Thus, this study further provides strong evidence that a severe semantic impairment is not always accompanied by surface dyslexia and that word reading can be selectively preserved although not relying on a semantically mediated pathway. These cases, in fact, provide further "evidence that non-semantic lexical processes are available to support oral word reading" (Raymer & Berndt, 1994, p. 481).

Overall, empirical evidence suggests that phonological dyslexia can occur in the context of a severe semantic impairment and is sometimes observed in cases that cannot be explained by the summation hypothesis. Instead, these data are in good agreement with the basic principles of the current model.

Is there empirical evidence that words can be accurately read via the sole semantic pathway? Two classes of models suggest that virtually all known words can be accurately read through the semantic pathway, namely PDP and multiple-level models. Among the proponents of multiple-level models, Manning and Warrington (1995; see also Orpwood & Warrington, 1995) provided empirical evidence in support of the claim that the semantic pathway is fully competent for real word reading. Two patients, JW and MRF, were described who showed a pattern of phonological dyslexia with intact word reading (98% and 96% correct for JW and MRF, respectively, without regularity or concreteness effects) and impaired pseudoword reading (about 50% correct for each participant) in the context of preserved comprehension. Their performance was viewed as reflecting the operation of an intact semantic route along with a damaged sublexical route.

This explanation in particular stems from the very significant effect of meaningfulness of letter combinations observed in pseudoword reading. Indeed, the performance of these patients in nonword reading increased with string familiarity; for instance, only 5% and 15% fewer familiar pseudowords (e.g., "xab") were read accurately against 55% and 60% more familiar ones (e.g., "bir"). In fact, it was observed that meaningfulness was highly correlated with N count so that a very high and significant correlation was found between the N values and the patients' scores in pseudoword reading. Furthermore, the two patients produced many lexicalization errors (>50%) that essentially occurred on pseudowords with a high N count. Manning and Warrington (1995) argued that better reading of the pseudowords that had a higher neighborhood size could be accounted for not by the residual capacities in the sublexical route but rather by the fact that reading of these pseudowords was entirely mediated by the lexico-semantic route.

Although it clearly appears that the effects reported by Manning and Warrington (1995) in pseudoword reading are not compatible with the operation of a so-called sublexical route, an account of these effects in terms of processing by a semantically mediated pathway seems rather dubious. As the links from orthography to meaning and from meaning to phonology are totally arbitrary, it is unclear how phonological representations

---

20 It could be argued that Italian requires less reliance on semantics in reading acquisition because of the quasi-regular mapping between orthography and phonology. This could explain why this patient remains able to read words despite a major semantic disorder but does not account for the word–nonword dissociation within the PMSF96 framework.
generated from pseudowords’ orthographic neighbors through
the semantic system might combine appropriately to give rise
to the expected phonological form of the target nonword (or to
a phonologically similar real word when an error occurs). The
pattern of performance showed by cases JW and RMF can no
more be accounted for within the single phonological pathway
postulated by PMSP96 because no deficit within this pathway
can result in a selective nonword reading impairment. In con-
trast, this pattern is fully compatible with a model that supposes
the existence of two reading modes as in the current one. Indeed,
the observed pattern of performance is predicted within this
model as resulting from reading in the global mode. It has
already been demonstrated that reading in global mode simulates
the pattern of phonological dyslexia. It is now shown that non-
word processing in the global reading mode also simulates the
pattern of performance shown by patients reported by Manning
and Warrington. Simulations conducted in the forced global
mode on 365 monosyllabic pseudowords from the original list
of 830 show that reading performance varies as a function of
$N$ count (see Figure 18).

Furthermore, as shown in Manning and Warrington’s (1995)
data, a very high correlation ($r = .94$) was found in the simula-
tions between the nonword reading scores and the $N$ count val-
ues. Finally, as already stressed in the simulation of phonological
dyslexia, and consistent with the patient data, lexicalization er-
rors were more prone to occur on those nonwords with a high
$N$ count (93.94%) than on the pseudowords without neighbors
(40%). Overall, the close correspondence between the simula-
tions and the performance of the two phonological dyslexic
patients, JW and MRF, reported by Manning and Warrington
supports the current model.

Other findings also argue against the hypothesis of an obliga-
tory semantic mediation in isolated word reading, in particular
the reading of very irregular words. Breen and Warrington
(1995; see also Breen & Warrington, 1994) reported the case of
a patient, NOR, who was able to read any kind of words,
including very irregular words, and nonwords despite exhibiting
a severe picture naming disorder (without any comprehension
deficit). This latter deficit suggests an impairment of the mean-
ing-to-phonology mapping procedures that should have been
accompanied, according to PMSP96 and in contrast to what
was empirically observed, by a selective deficit in irregular word
reading (i.e., a surface dyslexia). In order to accommodate these
data, Breen and Warrington (1995) claimed that “it is difficult
to account for these results in terms of a shared route to phonol-
ogy for reading and naming” (p. 587) and concluded “that
there must be more than one route to phonology” (p. 587).
Orpwood and Warrington (1995) explicitly claimed that there
must be “separate oral output stores for reading and naming”
(p. 255). This explanation in terms of multiple redundant pho-
nological lexicons does not appear very parsimonious. A more
natural account of these data would be to postulate the existence
of a fully competent phonological pathway for reading so that
performance could be dissociated in reading and picture naming.
That is precisely the conception of the reading system that is
advocated here.

**Developmental Dyslexia**

A number of findings clearly demonstrate the existence of
different varieties of developmental dyslexia (Boder, 1973; Cas-
tles & Coltheart, 1993; Manis, Seidenberg, Doi, McBride-
Chang, & Petersen, 1996; Marshall, 1984; Mattis, French,
& Rapin, 1975; Mitteer, 1982; Seymour & Mcgregor, 1984; Val-
dois, 1996b; Valdois, Gérard, Vanault & Dugas, 1995). Two
main subtypes have been identified that are developmental ana-
logues of the acquired forms of surface and phonological dys-
lexia. As in the acquired form, developmental surface dyslexia
is primarily characterized by a selective impairment in irregular
word reading, resulting in the production of regularization er-
ers, whereas a selective difficulty in pseudoword reading is
the defining feature of developmental phonological dyslexia.
However, the nature of the cognitive dysfunction(s) responsible
for these two distinct patterns of developmental dyslexia remains
a matter of ongoing debate. A phonological impairment has been
repeatedly found in association with developmental phonologi-
cal dyslexia, and a causal relation between poor phonological
awareness and difficulties in pseudoword reading has been hy-
pothesized (Broom & Doctor, 1995; Campbell & Butterworth,
1985; Howard & Best, 1996; Sartori & Job, 1982; Snowling &
Hulme, 1989; Temple & Marshall, 1983). In contrast, no reliable
hypothesis about the origin of developmental surface dyslexia
has been proposed so far. Most studies agree on at least two
points: first, that individuals with developmental surface dys-
lexia have failed to develop specific knowledge about word
spelling and, second, that these individuals have no problems
with tasks of phonological awareness, suggesting the absence of
an underlying phonological deficit (Castles & Coltheart, 1996;
Goulandris & Snowling, 1991; Hanley, Hastie, & Kay, 1992;
Manis et al., 1996). However, studies of developmental surface
dyslexia remain rather silent on the cause of this reading acquisi-
tion disorder. Goulandris and Snowling found that their partici-
pant’s surface dyslexia was associated with a poor performance
on tasks of visual memory; they hypothesized that this visual
memory deficit could be the cause of his poor acquisition of
orthographic knowledge. However, a number of other cases have
been reported of individuals with developmental surface dys-
exia who demonstrated normal visual memory abilities (Cas-
tles & Coltheart, 1996; Hanley et al., 1992; Romani & Stringer,
1994; Valdois, 1996a). The question of whether visual memory problems can sometimes result in the pattern of developmental surface dyslexia remains open as does the question of the origin of at least those forms that exhibit no associated memory problems.

Within the theoretical framework of the dual-route model, developmental surface and phonological dyslexias could arise following selective difficulties in acquiring either a lexical or sublexical procedure for reading (Castles & Coltheart, 1993; Coltheart et al., 1993). Unfortunately, the dual-route model offers no explanation of why these procedures would not develop normally in dyslexic children. In reference to the SM89 model, Manis et al., (1996) suggested that developmental surface dyslexia might arise from limitations in the available computational resources within the phonological route, whereas developmental phonological dyslexia was viewed as resulting from a deficit in phonological representation. The former impairment was simulated by PMSP96 in training a feedforward network with a limited number of hidden units (30 instead of 100). The authors demonstrated that limiting the number of hidden units selectively impaired performance on low-frequency irregular words whereas high-frequency irregular words, regular words, and pseudowords remained accurately processed. In fact, the simulation proposed by PMSP96 only approximated the pattern of developmental surface dyslexia because a strict dissociation between perfect high-frequency irregular word reading and poor low-frequency irregular word reading (less than 60% correct) has not been empirically reported; it seems rather that developmental surface dyslexics typically have a strong impairment that affects both high- and low-frequency irregular word reading (Castles & Coltheart, 1996; Valdois, 1996a; Valdois, Carbonnel, David, Ans, & Pellat, 1998). Furthermore, the pattern of developmental surface dyslexia was simulated in PMSP96 by “lesioning” a network (30 hidden units instead of 100) that had learned the orthographic–phonological correspondences of words without support from semantics. Because the authors assumed that activation from semantics in particular contributes to irregular word reading, it is probable that a simulation conducted in a network (with limited resources) trained with semantic support would affect irregular word reading even less.

PMSP96 further suggested that developmental phonological dyslexia could arise from the use of poor phonological representations, arguing that the SM89 network that employed inappropriate phonological and orthographic representations in some way behaved like an individual with phonological dyslexia with a performance over 97% correct in word reading and only 75% correct in pseudoword reading. However, SM89’s model at best only approximates the pattern of mild developmental phonological dyslexia and does not provide an account of the most severe forms of this disorder. Furthermore, the model provides no insight into the nature of the relationship between phonological dyslexia and phonemic awareness or between phonological dyslexia and short-term memory deficits (Snowling & Hulme, 1994).

The current model provides a general account of the two forms of surface and phonological developmental dyslexia and gives new insights on the kinds of deficits that could be at the origin of these developmental disorders. It first predicts, in a very general way, that some children might have great difficulty creating word traces even though they acquire segment traces in a relatively normal way. These children would then present a selective impairment on high- and low-frequency irregular words, thus showing the pattern characteristic of developmental surface dyslexia. Other children might have a selective difficulty in the acquisition of segment traces, thus leading to a pattern consistent with developmental phonological dyslexia. With respect to developmental surface dyslexia, an inability to create word traces would follow from a reduction of the visuo-attentional window through which information from the orthographic input is extracted. In fact, the interpretation of this developmental disorder is very similar to that proposed to account for acquired nonfluent surface dyslexia. It has been demonstrated previously that such a deficit resulted in the production of regularization errors and was sometimes accompanied by additional errors in the pronunciation of contextual graphemes (in particular: ‘c’, ‘g’, ‘s’); these two kinds of errors are found in developmental surface dyslexia. Furthermore, our interpretation is well in accordance with a recent case study of developmental surface dyslexia (Valdois, 1996a; Valdois et al., 1998) showing that this disorder was accompanied by a visual attentional deficit and that a remediation program essentially based on visual and visuo-attentional training considerably improved the reading performance of this child with dyslexia. Although developmental surface dyslexia is viewed here as resulting from a specific visuo-attentional impairment, it is nevertheless expected that the performance of individuals with surface dyslexia will be very similar to that of younger readers without dyslexia (matched for reading level) who have not yet established specific knowledge about familiar words. Overall, our account clearly supports a view of developmental surface dyslexia as a specific deficit pattern and not as a developmental delay pattern in contrast to Manis et al.'s (1996) interpretation.

With respect to developmental phonological dyslexia, at least two causes could be found for this disorder. It is first quite obvious that if the current network were trained only with whole word correspondences, its performance on pseudowords would be rather poor. Accordingly, it has been demonstrated that children who learned to read using a strict global method developed only very limited abilities to generalize (Beech, 1987). However, developmental phonological dyslexia could also result from an impairment at the phonological buffer level. As previously argued to account for acquired phonological dyslexia, a deficit of the blending process (at the phonological buffer level) would result in a selective difficulty in pronouncing the letter strings that are processed analytically. Such a deficit could also easily account for the fact that individuals with developmental phonological dyslexia show poor performance in metaphonological tasks that also involve the phonological buffer and that their deficit is frequently associated with working memory problems. A deficit of the blending process further predicts that the new words that are met without feedback and typically read analytically would never create a corresponding word trace, although the printed words processed in the global mode would normally lead to the creation of specific word traces. Given that low-frequency (regular and irregular) words are more likely to be new words for a child, it might be expected that a greater number of them would be read analytically by individuals with phonological dyslexia as compared with individuals without
phonological dyslexia (matched for chronological age). Consequently, the ability to read low-frequency irregular words might be limited in individuals with developmental phonological dyslexia as empirically reported in some cases.

**Visual Lexical Decision**

According to Coltheart et al. (1993), the decision as to whether an orthographic letter string is a word or not is made within the dual-route model by consulting the lexicon. The input letter string is recognized as a real word if its representation is found within the lexicon; it is rejected otherwise. Despite its apparent simplicity, this rather vague explanation is not at all sufficient to account for the whole complexity of human performance in lexical-decision tasks. Other evidence suggests that in this task, the decision can be taken at different levels of processing (Balota & Chumbley, 1984; James, 1975; Seidenberg, 1989, Seidenberg & McClelland, 1989) and that the locus of decision differs depending on the nature of the stimuli that are used in experiments. Moreover, it is probable that results from lexical-decision studies not only depend on the structural properties of the stimuli but also reflect the participant's strategy and hence "aspects of the decision process rather than normal processes in recognition" (Seidenberg, 1989, p. 69). Decision is probably primarily taken at the orthographic level when usual real words are mixed with illegal (unpronounceable) pseudowords or with pronounceable pseudowords made of unusual letter strings. The task can also involve more complex processing, taking into account not only the orthographic but also the phonological characteristics of the input strings, when lists contain very word-like pseudowords, strange real words, or both. Waters and Seidenberg (1985) indeed showed that no regularity effect was found on lists containing regular and irregular words and pronounceable pseudowords whereas a regularity effect was found on lists containing regular and irregular pseudowords or with pronounceable pseudowords made of unusual letter strings. The task can also involve more complex processing, taking into account not only the orthographic but also the phonological characteristics of the input strings, when lists contain very word-like pseudowords, strange real words, or both. Waters and Seidenberg (1985) indeed showed that no regularity effect was found on lists containing regular and irregular words and pronounceable pseudowords whereas a regularity effect appeared for low-frequency words when strange words were added to the list, suggesting that participants based their decision at the phonological level in the latter case. Hence, when the discrimination of words (including strange words) and pseudowords is difficult on the basis of orthography, participants use phonological information. Semantic processing can also be involved when words and very word-like pseudowords have to be discriminated as suggested by findings showing a concreteness effect in lexical decision (James, 1975). Dual-route models that typically assume that lexical decision is made by consulting the orthographic lexicon could only account for those results in which decision is obviously taken at the orthographic level, but an explanation of findings such as those of Waters and Seidenberg or James would be rather puzzling in this framework.

SM89's model can simulate lexical decision at the orthographic level; responses are based on the orthographic error score, which gives some measure of the orthographic familiarity of the input letter string. Their PDP model was, however, much poorer than skilled readers at lexical decision (Besner et al., 1990; Fera & Besner, 1992). SM89 claimed that they could potentially account for the results in which decision is taken at a phonological level in evaluating the familiarity of phonological outputs (or at the semantic level; Plaut, 1997). In the present model, lexical decision, when taken at an orthographic level, should be viewed as involving an essential component of the reading model, namely the orthographic checking process. An input pattern would be considered to be a word or not depending on the ability of the orthographic checking process to detect a difference between the O1 activity pattern generated from the orthographic input and the O2 activity pattern corresponding to the orthographic echo generated from the EM word-trace activity. Obviously, our model should perfectly simulate results from a lexical-decision task including orthographically unfamiliar pseudowords. Unlike standard views of lexical decision but in accordance with SM89, it was again demonstrated here that accounting for this task does not require access to individualized word representations because comparison was made between the orthographic input and an O2 orthographic echo resulting from the contribution of a number of EM word traces. An account of the results in which decision is taken at the phonological or semantic level is beyond the scope of the present model because no system has been currently implemented for evaluating the phonological or semantic familiarity of stimuli.

**Perspectives for Extending the Model**

In the current form of the network, the semantic pathway has not been implemented. However, this does not represent a true limitation because, as previously pointed out, we do not need to resort to the unimplemented semantic pathway to account for a number of important data, in particular the neuropsychological data, as SM89 and PMSP96 did. We rather claim that it is unreasonable to hypothesize that accurate isolated word reading essentially relies on semantic mediation. Furthermore, the present model offers some interesting perspectives on the way the semantic pathway could be implemented. The model assumes, in line with Hintzman (1986), that the EM layer keeps a specific trace of each new episode that minimally includes two components: a component corresponding to the orthographic event and the other to its environmental context (including its semantic context). This latter component has not yet been implemented, but its theoretical role already provides a natural account of the phenomena of frequency compression in the number of traces allocated to each word. The theoretical role of this component could also account for contextual effects in reading. Indeed, if the context associated with each occurrence of a word had been implemented, then the input to be read should have included two components, the orthographic input itself (as was the case) and its current context (in particular, the semantic context). This compound input would then better match those EM traces sharing a similar context. This would provide a natural account of some well-known experimental data from semantic priming tasks in which the congruous or incongruous nature of the prime (or context) is manipulated (Tulving & Gold, 1963). In the congruous condition, the input (word plus context) would strongly match the relevant EM traces (corresponding to the input word) that would have been learned in a context similar to the current one. Being then more strongly activated than in the neutral condition, these relevant EM traces might contribute to relaying stronger activation to the P layer units so that shorter naming latencies would be expected in the condition of congruous context. Conversely, in the incongruous context condition, matching with relevant EM traces will be comparable to that
observed in the neutral condition, but those irrelevant traces (corresponding to other words) sharing the current context will be more strongly activated than in the neutral condition, leading to longer naming latencies. It has been previously argued (Carbonnel et al., 1997; Damasio, 1989; Hintzman, 1986; Rousset & Schreiber, 1992; Schreiber, Rousset, & Tiberghien, 1991) that the meaning of an entity was not abstracted and permanently stored but was rather temporarily recreated when the entity is present in a current episode and arises from the reactivation of previous episodes. In such a framework, the currently evoked meaning is held to result from the combined contents of those episodic traces activated by the current entity. The implementation of the contextual component of EM traces would then amount to a simulation of meaning evocation viewed in this way. The EM traces differentially activated by a given input word would trigger the activation of their contextual components, the meaning evoked by the orthographic input being then the combined content of these different components. Adaptive connections between the contextual component and the phonological layer would lead to complete the implementation of the orthography-to-meaning-to-phonology pathway.

**Conclusion**

The model we have described supports a view of the reading system according to which reading relies on two successive procedures, a global procedure using knowledge about whole word correspondences and an analytic procedure based on the activation of word syllabic segments. These procedures are not dedicated to a specific kind of item, but rather apply as a function of the system's ability to recognize the input as familiar or unfamiliar. As argued by the dual-route model, it is probable that the theoretical choice of two reading procedures was determinant in the capacity of the current model to account for the basic features of both normal and pathological performance; it was also a critical determinant of the network's ability to process polysyllabic words and pseudowords as well as monosyllabic items. On the other hand, and more in accordance with the PDP connectionist approach (SM89; PMSP96), the two procedures do not rely on radically different computational principles, pronunciation being always computed from knowledge about previously experienced exemplars. It was demonstrated that this view of the reading system contributes to the understanding of a wide range of normal and pathological data and goes beyond previous models in accounting for the reading of polysyllabic items.

**References**


Hanley, R., Hastie, K., & Kay, J. (1992). Developmental surface dys-
Lexical analogies do the work of non-lexical rules.


Vincent, S. B. (1912). The function of the vibrissae in the behavior of the white rat. Behavioral Monographs, 1, (Serial No. 5).


Appendix A

Orthography-to-Phonology Correspondences for Isolated Consonantal and Vocalic Graphemes

| b  | /b*/ | x  | /ks*/ | e  | /e/ | o  | /u/ |
| c  | /k*/ | z  | /z*/ | ei | /e/ | ou | /u/ |
| d  | /d*/ | ph | /ph/ | eu | /e/ | oi | /u/ |
| f  | /f*/ | ch | /ch/ | ed | /e/ | oã | /ã/ |
| g  | /g*/ | qu | /kw/ | en | /e/ | on | /o/ |
| h  | /h*/ | a  | /a/ | eau | /o/ | ou | /u/ |
| jk | /j*/ | à  | /a/ | i  | /i/ | ouin | /ui/ |
| l  | /l*/ | ã  | /ã/ | t  | /t/ | u  | /y/ |
| m  | /m*/ | a  | /æ/ | y  | /i/ | õ  | /y/ |
| n  | /n*/ | al | /æ/ | ir | /r/ | õ | /y/ |
| p  | /p*/ | au | /au/ | in | /i/ | un | /u/ |
| q  | /q*/ | an | /æ/ | o  | /o/ | ui | /y/ |
| r  | /r*/ | ain | /æ/ | o  | /o/ | uf | /y/ |
| s  | /s*/ | e  | /æ/ | õ | /o/ | uin | /y/ |
| t  | /t*/ | é  | /e/ | oe | /e/ | y  | /i/ |
| v  | /v*/ | è  | /e/ | oeu | /e/ | yn | /y/ |
| w  | /w*/ | ã  | /ã/ | oi | /u/ | ay | /e/ |
| ey | /e/ | oy | /ia/ | uy | /y/ |

Note. /* denotes a silent vocalic phoneme.

Appendix B

Errors in Global Reading of Words Leading to the Production of Another Word (With Their Frequency)

<table>
<thead>
<tr>
<th>Word</th>
<th>Frequency</th>
<th>Word</th>
<th>Frequency</th>
<th>Word</th>
<th>Frequency</th>
<th>Word</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>résident</td>
<td>110</td>
<td>président</td>
<td>8,049</td>
<td>plume</td>
<td>5,105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>caillot</td>
<td>119</td>
<td>caillou</td>
<td>2,156</td>
<td>lapin</td>
<td>2,102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dégelier</td>
<td>123</td>
<td>dèceler</td>
<td>1,127</td>
<td>vengeance</td>
<td>2,122</td>
<td></td>
<td></td>
</tr>
<tr>
<td>claire</td>
<td>127</td>
<td>clair</td>
<td>13,741</td>
<td>violet</td>
<td>974</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rétention</td>
<td>127</td>
<td>prétention</td>
<td>2,284</td>
<td>tribu</td>
<td>1,340</td>
<td></td>
<td></td>
</tr>
<tr>
<td>affect</td>
<td>136</td>
<td>affecter</td>
<td>2,544</td>
<td>osé</td>
<td>2,654</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vitreux</td>
<td>144</td>
<td>vitre</td>
<td>3,616</td>
<td>bouton</td>
<td>2,931</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cuisson</td>
<td>148</td>
<td>buisson</td>
<td>1,591</td>
<td>repentir</td>
<td>1,114</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ponctuer</td>
<td>157</td>
<td>ponctuel</td>
<td>310</td>
<td>compléter</td>
<td>1,748</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sis</td>
<td>161</td>
<td>six</td>
<td>14,018</td>
<td>portière</td>
<td>1,220</td>
<td></td>
<td></td>
</tr>
<tr>
<td>coursier</td>
<td>170</td>
<td>courrier</td>
<td>2,424</td>
<td>grille</td>
<td>2,926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gemme</td>
<td>174</td>
<td>germe</td>
<td>2,288</td>
<td>refaire</td>
<td>2,744</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cuisiner</td>
<td>182</td>
<td>cuisine</td>
<td>7,287</td>
<td>fumeux</td>
<td>365</td>
<td></td>
<td></td>
</tr>
<tr>
<td>stable</td>
<td>195</td>
<td>stable</td>
<td>991</td>
<td>masse</td>
<td>9,572</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hermine</td>
<td>233</td>
<td>vermine</td>
<td>446</td>
<td>canot</td>
<td>684</td>
<td></td>
<td></td>
</tr>
<tr>
<td>empreinte</td>
<td>276</td>
<td>empreinte</td>
<td>1,332</td>
<td>fonction</td>
<td>10,376</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pesse</td>
<td>300</td>
<td>pèse</td>
<td>3,390</td>
<td>bénir</td>
<td>1,182</td>
<td></td>
<td></td>
</tr>
<tr>
<td>terref</td>
<td>353</td>
<td>terreur</td>
<td>4,475</td>
<td>délice</td>
<td>2,190</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hériter</td>
<td>395</td>
<td>hérétique</td>
<td>4,573</td>
<td>éclair</td>
<td>923</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aisselle</td>
<td>446</td>
<td>vaisselle</td>
<td>1,038</td>
<td>perpétuer</td>
<td>3,607</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pouls</td>
<td>544</td>
<td>poule</td>
<td>2,182</td>
<td>bouillir</td>
<td>872</td>
<td></td>
<td></td>
</tr>
<tr>
<td>saillie</td>
<td>625</td>
<td>saillir</td>
<td>323</td>
<td>heure</td>
<td>86,928</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bile</td>
<td>710</td>
<td>bile</td>
<td>285</td>
<td>ingénieur</td>
<td>1,586</td>
<td></td>
<td></td>
</tr>
<tr>
<td>habiter</td>
<td>1,357</td>
<td>habituel</td>
<td>3,773</td>
<td>officier</td>
<td>3,335</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Appendices continue)
Appendix C

Errors in Global Reading of Pseudowords Leading to Lexicalizations

<table>
<thead>
<tr>
<th>Target</th>
<th>Response</th>
<th>Target</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>bagle</td>
<td>balle</td>
<td>balage</td>
<td>bagage</td>
</tr>
<tr>
<td>blot</td>
<td>bloc</td>
<td>botton</td>
<td>båton</td>
</tr>
<tr>
<td>bourgon</td>
<td>bourdon</td>
<td>bousse</td>
<td>bourse</td>
</tr>
<tr>
<td>bran</td>
<td>brun</td>
<td>chuse</td>
<td>chose</td>
</tr>
<tr>
<td>clon</td>
<td>clou</td>
<td>colde</td>
<td>corde</td>
</tr>
<tr>
<td>crate</td>
<td>crête</td>
<td>cron</td>
<td>crin</td>
</tr>
<tr>
<td>fiance</td>
<td>finance</td>
<td>laupe</td>
<td>lampe</td>
</tr>
<tr>
<td>laine</td>
<td>moine</td>
<td>méduse</td>
<td>méduse</td>
</tr>
<tr>
<td>puis</td>
<td>puvis</td>
<td>pérat</td>
<td>pérl</td>
</tr>
<tr>
<td>pertuche</td>
<td>pertuche</td>
<td>peute</td>
<td>peute</td>
</tr>
<tr>
<td>poncher</td>
<td>pencher</td>
<td>prose</td>
<td>prose</td>
</tr>
<tr>
<td>puine</td>
<td>rüne</td>
<td>roison</td>
<td>poison</td>
</tr>
<tr>
<td>sabane</td>
<td>cabane</td>
<td>stal</td>
<td>star</td>
</tr>
<tr>
<td>stort</td>
<td>store</td>
<td>stule</td>
<td>style</td>
</tr>
<tr>
<td>tacle</td>
<td>table</td>
<td>taure</td>
<td>terre</td>
</tr>
<tr>
<td>tinne</td>
<td>tonne</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Appendix D

Output Activity of Episodic Memory Layer EM, of the Orthographic layer O2, and of the Phonological Layer P

When the Network is Reading in Global Mode the Pseudoword *virdin* (*viRdë*)

Layer EM Activity

In Table D1, the first two columns contain a number of EM word traces activated (among 778 active traces) by the input pseudoword *virdin* and their associated frequency factor \( \varphi \) (with \( 0 < \varphi < 1 \)). The third column gives the orthographic similarity \( a_i \) (with \( 0 \leq a_i \leq 1 \)) between the tested pseudoword and the word traces. The highest output values \( s_i \) of the corresponding EM macro units, resulting from Equation 9, are listed in the fourth column (in descending order and rounded to three decimal places). The input activation of the orthographic layer O2 and of the phonological layer P will result from the parallel summation of all learned orthographic words (and all phonological correspondences, respectively), each contributing in proportion to the corresponding EM output value.

Table D1

Layer EM Activity for the Input Pseudoword "Virdin"

<table>
<thead>
<tr>
<th>EM word traces</th>
<th>EM matching with #virdin#</th>
<th>Phonological correspondences stored in layer P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learned words</td>
<td>( \varphi )</td>
<td>Similarity</td>
</tr>
<tr>
<td>#jardin#</td>
<td>1.000</td>
<td>0.750</td>
</tr>
<tr>
<td>#vilain#</td>
<td>0.881</td>
<td>0.750</td>
</tr>
<tr>
<td>#airain#</td>
<td>0.231</td>
<td>0.750</td>
</tr>
<tr>
<td>#verdin#</td>
<td>0.104</td>
<td>0.625</td>
</tr>
<tr>
<td>#vision#</td>
<td>0.998</td>
<td>0.625</td>
</tr>
<tr>
<td>#pardon#</td>
<td>0.993</td>
<td>0.625</td>
</tr>
<tr>
<td>#miroir#</td>
<td>0.981</td>
<td>0.625</td>
</tr>
<tr>
<td>#tiroir#</td>
<td>0.873</td>
<td>0.625</td>
</tr>
<tr>
<td>#violon#</td>
<td>0.750</td>
<td>0.625</td>
</tr>
<tr>
<td>#cardif#</td>
<td>0.691</td>
<td>0.625</td>
</tr>
<tr>
<td>#serein#</td>
<td>0.597</td>
<td>0.625</td>
</tr>
<tr>
<td>#air#</td>
<td>1.000</td>
<td>0.600</td>
</tr>
<tr>
<td>#vie#</td>
<td>1.000</td>
<td>0.600</td>
</tr>
<tr>
<td>#vif#</td>
<td>1.000</td>
<td>0.600</td>
</tr>
<tr>
<td>#cordon#</td>
<td>0.556</td>
<td>0.625</td>
</tr>
</tbody>
</table>
When testing the pseudoword *virdin*, the resulting input activations of the letter units belonging to each positional WTA cluster are given in Table D2 (only the three main nonzero activations are given within each O2 cluster). Letter units can have an initial effective activity only if their input activation is equal to or above the orthographic threshold $e_0 = 0.35$. Within each cluster, only the above-threshold letter unit (boldfaced values in the table) will have a final stabilized output activity equal to one, whereas the other units will not be active at all. It can be seen that the orthographic echo (the sequence of boldfaced letters), which has emerged from the reverberation of the orthographic input on the EM word traces, has exactly the same orthographic form as the pseudoword presented in test: In this case, the input orthographic object is said to be recognized by the network.

Table D2
Layer O2 Activity for the Input Pseudoword Virdin

<table>
<thead>
<tr>
<th>Cluster -2</th>
<th>Cluster -1</th>
<th>Cluster 0</th>
<th>Cluster +1</th>
<th>Cluster +2</th>
<th>Cluster +3</th>
<th>Cluster +4</th>
<th>Cluster +5</th>
<th>Cluster +6</th>
<th>Cluster +7</th>
</tr>
</thead>
<tbody>
<tr>
<td>#: 0.999</td>
<td>v: 0.425</td>
<td>li: 0.514</td>
<td>r: 0.615</td>
<td>d: 0.371</td>
<td>li: 0.739</td>
<td>m: 0.690</td>
<td>#: 0.838</td>
<td>#: 0.017</td>
<td>#: 0.005</td>
</tr>
<tr>
<td>j: 0.250</td>
<td>a: 0.339</td>
<td>l: 0.237</td>
<td>a: 0.312</td>
<td>o: 0.990</td>
<td>r: 0.092</td>
<td>e: 0.014</td>
<td>l: 0.004</td>
<td>&amp;: 0.001</td>
<td></td>
</tr>
<tr>
<td>a: 0.072</td>
<td>e: 0.088</td>
<td>s: 0.039</td>
<td>#: 0.074</td>
<td>#: 0.024</td>
<td>#: 0.034</td>
<td>a: 0.005</td>
<td>t: 0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Layer P Activity
In response to the input pseudoword *virdin*, the resulting input activations of the phoneme units belonging to each positional WTA cluster of the relevant syllable are given in Table D3 (only the three main nonzero activations are given within each P cluster). Phoneme units can have an initial effective activity only if their input activation is equal to or above the phonological threshold $e_p = 0.17$. From these above-threshold activations (values in boldface in the table), the phoneme units eventually compete within each cluster where only the winning unit (underlined values) will have a final stabilized output activity equal to one, whereas the activity of the other units will tend toward zero. It can be seen that the phonological echo (the sequence of underlined phonemes), which has emerged from the reverberation of the orthographic input on the EM word traces, corresponds to the expected pronunciation (/viRd&W/) of the pseudoword presented in the test.

Table D3
Layer P Activity for the Input Pseudoword Virdin

<table>
<thead>
<tr>
<th>Syllables</th>
<th>Cluster -2</th>
<th>Cluster -1</th>
<th>Cluster 0</th>
<th>Cluster +1</th>
<th>Cluster +2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syllable 1</td>
<td>/v/: 0.028</td>
<td>/v/: 0.398</td>
<td>/l/: 0.409</td>
<td>/l/: 0.454</td>
<td></td>
</tr>
<tr>
<td></td>
<td>/l/: 0.003</td>
<td>/l/: 0.251</td>
<td>/a/: 0.345</td>
<td>/l/: 0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>/p/: 0.002</td>
<td>/l/: 0.056</td>
<td>/l/: 0.143</td>
<td>/l/: 0.010</td>
<td></td>
</tr>
<tr>
<td>Syllable 2</td>
<td>/R/: 0.045</td>
<td>/d/: 0.364</td>
<td>/R/: 0.579</td>
<td>/R/: 0.082</td>
<td>/h/: 0.003</td>
</tr>
<tr>
<td></td>
<td>/z/: 0.025</td>
<td>/l/: 0.246</td>
<td>/l/: 0.115</td>
<td>/l/: 0.018</td>
<td></td>
</tr>
<tr>
<td></td>
<td>/z/: 0.006</td>
<td>/R/: 0.128</td>
<td>/R/: 0.090</td>
<td>/l/: 0.007</td>
<td></td>
</tr>
<tr>
<td>Syllable 3</td>
<td>/a/: 0.019</td>
<td>/a/: 0.027</td>
<td>/l/: 0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/j/: 0.008</td>
<td>/a/: 0.004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/h/: 0.003</td>
<td>/a/: 0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syllable 4</td>
<td>/h/: 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Boldface letters and values represent phonemes with an above-threshold activation. Boldfaced and underlined characters denote the “winning” phonemes within the corresponding clusters.

(Appendices continue)
Appendix E

Output Activity of Episodic Memory Layer EM and of the Phonological Layer P When the Network is Reading in Global Mode a Low-Frequency Irregular Word With Frequent Enemies and Without Enemies
(Notations Are the Same as in Appendix D)

An Example With an Input Irregular Word With Frequent Enemies: gin

Layer EM activity. By classical definition, the word gin (/dʒi:n/) is irregular because it includes the letter string ‘in’ that is exceptionally pronounced /n/; indeed, this grapheme is pronounced /n/ in most other French words (called antagonistic words in the text), whatever the grapheme position. In Table E1, it can be seen that within the set of all possible antagonistic words for ‘gin’, only the high-frequency neighbors have their EM word traces substantially activated—in fact, the three high-frequency enemies /m/‘ne/, /v/‘n/, /j/‘n/. In contrast, an antagonist word trace, like engin=/dʒi:n/, will match only very weakly with gin and will therefore have no effect on P-layer activity.

Table E1
Layer EM Activity for the Input Word Gin

<table>
<thead>
<tr>
<th>EM word traces</th>
<th>EM matching with #gin#</th>
<th>Phonological correspondences stored in layer P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learned words</td>
<td>ϕ</td>
<td>Similarity</td>
</tr>
<tr>
<td># gin #</td>
<td>.314</td>
<td>1.000</td>
</tr>
<tr>
<td># fin #</td>
<td>1.000</td>
<td>.800</td>
</tr>
<tr>
<td># vin #</td>
<td>1.000</td>
<td>.800</td>
</tr>
<tr>
<td># pian #</td>
<td>.916</td>
<td>.800</td>
</tr>
<tr>
<td># lian #</td>
<td>.469</td>
<td>.800</td>
</tr>
<tr>
<td># bon #</td>
<td>1.000</td>
<td>.600</td>
</tr>
<tr>
<td># mon #</td>
<td>1.000</td>
<td>.600</td>
</tr>
<tr>
<td># son #</td>
<td>1.000</td>
<td>.600</td>
</tr>
<tr>
<td># ton #</td>
<td>1.000</td>
<td>.600</td>
</tr>
<tr>
<td># don #</td>
<td>.999</td>
<td>.600</td>
</tr>
<tr>
<td># grin #</td>
<td>.987</td>
<td>.600</td>
</tr>
<tr>
<td># gai #</td>
<td>.969</td>
<td>.600</td>
</tr>
<tr>
<td># ben #</td>
<td>.954</td>
<td>.600</td>
</tr>
<tr>
<td># pan #</td>
<td>.869</td>
<td>.600</td>
</tr>
<tr>
<td># gn #</td>
<td>.238</td>
<td>.600</td>
</tr>
<tr>
<td># ban #</td>
<td>.215</td>
<td>.600</td>
</tr>
<tr>
<td># gui #</td>
<td>.152</td>
<td>.600</td>
</tr>
</tbody>
</table>

Layer P activity. The above EM word traces corresponding to high-frequency enemies are sufficiently activated to contribute significantly to phonemic activation. It can be seen in Table E2 that competition occurs within the 0 cluster (between the “irregular” target phoneme /i/ and the “regular” phoneme /b/), increasing the time needed for cleaning up the cluster and, therefore, the whole network latency that is determined by this cluster alone.

Table E2
Layer P Activity for the Input Word Gin

<table>
<thead>
<tr>
<th>Syllables</th>
<th>Cluster −2</th>
<th>Cluster −1</th>
<th>Cluster 0</th>
<th>Cluster +1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syllable 1</td>
<td>/b/: 0.607</td>
<td>/b/: 0.608</td>
<td>/b/: 0.608</td>
<td>/b/: 0.607</td>
</tr>
<tr>
<td></td>
<td>/g/: 0.003</td>
<td>/l/: 0.107</td>
<td>/l/: 0.366</td>
<td>/l/: 0.016</td>
</tr>
</tbody>
</table>

Note. Boldfaced letters and values represent phonemes with an above-threshold activation. Boldfaced and underlined characters denote the “winning” phonemes within the corresponding clusters.

An Example With an Input Irregular Word Without Enemies: Chorus

Layer EM activity. The word chorus (/kɔrəs/) would, by classical definition, be irregular because it includes the letter string ch-‘k/ç/, which is pronounced /ʃ/ in most other French words. When testing chorus, it can be seen in Table E3 that the antagonistic EM word traces are weakly activated because they share only a few orthographic features with the target (e.g., cher−væn).
Table E3
Layer EM Activity for the Input Word Chorus

<table>
<thead>
<tr>
<th>EM word traces</th>
<th>Learned words</th>
<th>(\varphi)</th>
<th>Similarity</th>
<th>EM output</th>
<th>Syllable 1</th>
<th>Syllable 2</th>
<th>Syllable 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>'chorus'</td>
<td>.216</td>
<td>1.000</td>
<td>.782</td>
<td>ko</td>
<td>Rys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'choral'</td>
<td>.292</td>
<td>.750</td>
<td>.025</td>
<td>ko</td>
<td>Ral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chor'</td>
<td>1.000</td>
<td>.667</td>
<td>.019</td>
<td>Sè R</td>
<td>Rys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'choir'</td>
<td>.977</td>
<td>.667</td>
<td>.018</td>
<td>Sè k</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chobus'</td>
<td>.932</td>
<td>.667</td>
<td>.017</td>
<td>o</td>
<td>by</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'char'</td>
<td>.904</td>
<td>.667</td>
<td>.017</td>
<td>Sè R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chacque'</td>
<td>1.000</td>
<td>.625</td>
<td>.008</td>
<td>Sè k</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chargé'</td>
<td>1.000</td>
<td>.625</td>
<td>.008</td>
<td>Sè R</td>
<td>jé</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'charme'</td>
<td>.997</td>
<td>.625</td>
<td>.008</td>
<td>kè s</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'charge'</td>
<td>.994</td>
<td>.625</td>
<td>.008</td>
<td>kè s</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'opus'</td>
<td>.991</td>
<td>.625</td>
<td>.008</td>
<td>Sè R</td>
<td>jé</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'charir'</td>
<td>.360</td>
<td>.667</td>
<td>.007</td>
<td>o</td>
<td>pys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'charrois'</td>
<td>.508</td>
<td>.625</td>
<td>.004</td>
<td>Sè R</td>
<td>Rys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'charnu'</td>
<td>.334</td>
<td>.625</td>
<td>.003</td>
<td>Sè k</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chèque'</td>
<td>.384</td>
<td>.625</td>
<td>.003</td>
<td>Sè k</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chaps'</td>
<td>.351</td>
<td>.625</td>
<td>.003</td>
<td>Sè R</td>
<td>tÀ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'blockus'</td>
<td>.321</td>
<td>.625</td>
<td>.003</td>
<td>blo</td>
<td>kys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'corpse'</td>
<td>1.000</td>
<td>.571</td>
<td>.003</td>
<td>Sè R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chose'</td>
<td>1.000</td>
<td>.571</td>
<td>.003</td>
<td>Sè R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'héros'</td>
<td>.997</td>
<td>.571</td>
<td>.003</td>
<td>'é R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chère'</td>
<td>.996</td>
<td>.571</td>
<td>.003</td>
<td>Sè R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chéri'</td>
<td>.987</td>
<td>.571</td>
<td>.002</td>
<td>Sè R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'cours'</td>
<td>.358</td>
<td>.571</td>
<td>.002</td>
<td>kù R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'choque'</td>
<td>.251</td>
<td>.625</td>
<td>.002</td>
<td>fò kÀ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chorax'</td>
<td>.194</td>
<td>.625</td>
<td>.002</td>
<td>tò Rák s</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chique'</td>
<td>.180</td>
<td>.625</td>
<td>.001</td>
<td>Sè kÀ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chaht'</td>
<td>.174</td>
<td>.625</td>
<td>.001</td>
<td>Sè y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'choriste'</td>
<td>.292</td>
<td>.600</td>
<td>.001</td>
<td>ko Rys</td>
<td>tÀ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'choquer'</td>
<td>.720</td>
<td>.556</td>
<td>.001</td>
<td>So kÀ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chique'</td>
<td>.145</td>
<td>.625</td>
<td>.001</td>
<td>Sè kÀ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'clores'</td>
<td>.392</td>
<td>.571</td>
<td>.001</td>
<td>kò RÀ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'corde'</td>
<td>.324</td>
<td>.571</td>
<td>.001</td>
<td>'ò R dÀ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'viru'</td>
<td>.273</td>
<td>.571</td>
<td>.001</td>
<td>vi Rys</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'chorale'</td>
<td>.393</td>
<td>.556</td>
<td>.001</td>
<td>ko RÀ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'humus'</td>
<td>.191</td>
<td>.571</td>
<td>.000</td>
<td>y mÁ</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Layer P activity. The antagonistic EM word traces of the word chorus, being weakly activated, contribute minimally to phonemic input activation. The above-threshold activation of the “irregular” target phoneme /kÀ/, within the positional cluster -1, is not hindered by the concurrent “regular” phoneme /SÈ/ as seen in Table E4. Thus, for a low-frequency irregular word without enemies, such as chorus, the time needed to reach the stabilized final state of the P layer is not expected to be significantly different from that required when testing a matched regular word, and therefore, it is no more likely that a significant difference in naming latency will be observed.

Table E4
Layer P Activity for the Input Word Chorus

<table>
<thead>
<tr>
<th>Syllables</th>
<th>Cluster -2</th>
<th>Cluster -1</th>
<th>Cluster 0</th>
<th>Cluster +1</th>
<th>Cluster +2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syllable 1</td>
<td>/À/: 0.004</td>
<td>/À/: 0.825</td>
<td>/À/: 0.871</td>
<td>/À/: 0.074</td>
<td>/À/: 0.018</td>
</tr>
<tr>
<td></td>
<td>/À/: 0.003</td>
<td>/À/: 0.128</td>
<td>/À/: 0.068</td>
<td>/À/: 0.255</td>
<td></td>
</tr>
<tr>
<td></td>
<td>/À/: 0.001</td>
<td>/À/: 0.006</td>
<td>/À/: 0.025</td>
<td>/À/: 0.029</td>
<td></td>
</tr>
<tr>
<td>Syllable 2</td>
<td>/À/: 0.829</td>
<td>/À/: 0.827</td>
<td>/À/: 0.795</td>
<td>/À/: 0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>/À/: 0.029</td>
<td>/À/: 0.046</td>
<td>/À/: 0.026</td>
<td>/À/: 0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>/À/: 0.018</td>
<td>/À/: 0.029</td>
<td>/À/: 0.005</td>
<td>/À/: 0.002</td>
<td></td>
</tr>
</tbody>
</table>

Note. Boldfaced letters and values represent phonemes with an above-threshold activation.