Neural-Networks Models of Cognition
J.W. Donahoe and V. Packard Dorsel (Editors)
© 1997 Elsevier Science B.V. All rights reserved.

CHAPTER 27

A RECURRENT-NETWORK ACCOUNT OF READING, SPELLING, AND DYSLEXIA

Guy C. Van Orden, Anna M.T. Bosman,
Stephen D. Goldinger, and William T. Farrar, IV
Arizona State University

ABSTRACT

We present a recurrent-network account of printed word perception, spelling, and dyslexia. Our account explains why phonology is fundamental to reading and spelling, and why spelling is more difficult than reading. It also provides a basis for simulating the behavior displayed by children with reading problems (developmental dyslexia) and by adults with acquired reading problems (acquired dyslexia).

Introduction

Recurrent networks are connectionist models in which activation flows from input to output and back again, creating feedback loops. *Behavior* is modeled in self-organizing patterns of activation, but activation in any part of the network is always reflected throughout the network. Bidirectional flow of activation binds activation at each part to activation at every other part. This holism is the basis of our claim to a *mind as embodied action* metaphor: Perception and action are emergent in the coupling (resonance) of an organism and its environment (Varela, Thompson, & Rosch, 1991). One entailment of this metaphor is that an organism and its environment are themselves interdependent (cf. Turvey & Carello, 1981). In this chapter, we illustrate these claims in an explanation of reading and spelling that assumes interdependence between readers and the printed word. The empirical basis of our account is performance in laboratory reading and spelling tasks.

The description presented in this chapter derives from a general theoretical framework proposed by us and our colleagues (Grossberg & Stone, 1986; Stone & Van Orden, 1994; Van Orden & Goldinger, 1994; Van Orden, Pennington, & Stone, 1990, 1996), and is rooted in mathematical dynamic-systems theory (cf. Thelen & Smith, 1994). Clarifying the metaphor of *mind as embodied action* will be easier once we have presented our account in more detail. We begin by discussing a classic cognitive phenomenon: phonologic mediation in reading and spelling. Then, we describe a resonance theory of word perception and spelling. At its heart, we offer an explanation of why phonology is fundamental to reading and spelling, and why spelling is more difficult than reading. This account predicts non-intuitive "feedback" phenomena that have been corroborated in laboratory studies. We describe these studies, and then suggest how a recurrent network might accommodate the anomalous reading behavior of developmental and acquired dyslexics. In the final section, we return to the metaphor that underlies our approach.

Phonologic Mediation in Reading and Spelling

The emphasis in much reading research is on the perception of single words, as this is the most important aspect of reading skill. Poor word perception severely limits the development of skilled reading and reading comprehension (Perfetti, 1985). A perennial question in such research is whether a word's phonology (loosely, the "sound" of a word) influences visual word perception. A recent accumulation of empirical findings forcefully suggests that phonology's role in word perception is fundamental. Numerous experiments have shown that the phonology of a letter string affects its perception in simple reading tasks. (For overviews, see Berent & Perfetti, 1995; Van Orden et al, 1990.) For example, subjects tend to mistake ROZE for an exemplar of FLOWER in a categorization task, or they misclassify the letter string SUTE as a word in a lexical decision task, or they overlook misspellings such as MUNKEY in a proofreading task (Van Orden, Stone, Garlington, Markson, Pinnt, Simonfy, & Brichetto, 1992). And in writing, systematic misspellings, such as substituting ROZE or ROWS for ROSE are common. (For an overview, see Bosman & Van Orden, in press.) These errors indicate that phonology is central to reading and spelling. With respect to phonology, a ROZE is a ROWS is a ROSE. Note, however, that phonology is not explicit in these printed forms; it is only implicit with respect to the knowledge that readers bring to reading and spelling.

The effects of phonology are apparent with readers and writers spanning the full range of reading skill (i.e., beginning, skilled, and disabled readers). They are found across languages (in both alphabetic and non-alphabetic writing systems) and across laboratory tasks. These tasks include naming (quickly reading words aloud), lexical decision (quickly classifying words versus nonwords), semantic categorization (quickly determining whether words belong to designated categories), and proofreading (carefully checking a document for spelling errors). Why is phonology so involved in reading or spelling? It clearly is not always helpful, often leading to errors in these experimental tasks. In the next section, we describe a recurrent-network account of word perception and spelling. Our account pertains to a very simple recurrent network that has been implemented (Farrar & Van Orden, 1994), but the principled basis of our account is not tied to the specifics of our simulation. No claim is made with respect to a "correct" architecture (see Stone & Van Orden, 1994; Van Orden & Goldinger, 1994; Van Orden et al, 1990, 1996). We return to this issue in the final section of the chapter.

Reading and Spelling are Fundamentally Related

Imagine a fictitious nervous system that perceives printed words. This system consists of three families of neurons: letter neurons, phoneme neurons, and semantic neurons. Every neuron in each family is (potentially) bidirectionally connected to every neuron of the other two families. Bidirectionally connected means that if a feedforward connection exists from neuron "x" to neuron "y," there is also a feedback connection from neuron "y" to "x." Now, imagine a specific pattern of activation across the letter neurons, due to the presentation of a printed word. This letter pattern feeds activation forward through a matrix of "synaptic" connections, creating patterns of activation across phoneme and semantic neurons. The phoneme and semantic neurons, in turn, feed activation back through a top-down matrix of connections, transforming their patterns back into letter patterns. Whenever the feedback patterns match the original letter pattern, top-down activation *conserves* bottom-up activation. Consequently, the "matched" letter neurons conserve their capacity to reactivate matching phoneme and semantic neurons that, in turn, reactivate the letter neurons, and so on. This feedback cycle is temporarily stable, resulting in a coherent dynamic whole: a *resonance*.

This simple neural network is only for exposition. It is helpful to consider word perception in terms of artificial neural activity, but analogies between cognitive systems and actual nervous systems, albeit compelling, are limited. We conceive of word perception in cognitive terms. **Figure 1** illustrates cognitive macrodynamics of word perception (Van Orden & Goldinger, 1994) and spelling (Bosman & Van Orden, in press), and **Figure 2** illustrates microdynamics.

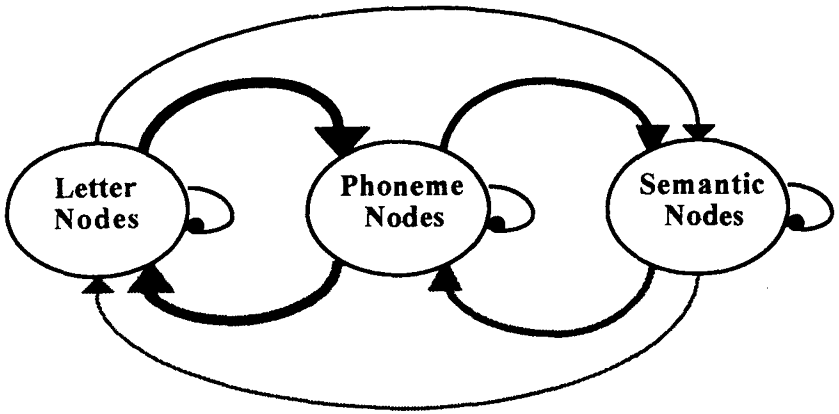

FIGURE 1. Macrodynamics of reading and spelling performance emergent in a recurrent network. The boldness of the arrows indicates the overall strength of the relations between letter, phoneme, and semantic node families (see text).

Figure 1 portrays a recurrent network with three families of fully interdependent nodes (letter nodes, phoneme nodes, and semantic nodes). On average, the connections between node families differ in strength; the rank order of overall strength is illustrated by the relative boldness of arrows in the figure. In alphabetic languages, letters and phonemes correlate quite strongly. For example, the letter B is almost always pronounced as /b/, and the phoneme /b/ is always written with a B. Correlations between phonemes and semantic features, or letters and semantic features, are far weaker than correlations between letters and phonemes. Knowing that a word begins with the letter B indicates almost nothing about its meaning, but much about its initial pronunciation.

Notice also that phoneme-semantic connections are depicted as stronger than letter-semantic connections, primarily because we speak before and more often than we read. Moreover, once in place, this asymmetry is self-perpetuating. Reading strengthens phoneme-semantic connections, because phonology functions in every instance of printed-word perception. Thus, even the exceptional condition of people who read more than they speak would support phoneme-semantic connections that are at least as strong as letter-semantic connections. Also, if a coherent positive-feedback loop forms from semantic to phoneme nodes before the feedback loop from semantic to letter nodes, then printed or spoken discourse may proceed without the contribution of the feedback loop from semantic to letter nodes. The absence of resonance in the latter feedback loop may preclude strengthening the connections between letter and semantic nodes (see discussion below, and Grossberg & Stone, 1986). Thus, at this macro-level of description, families of nodes differ in the overall strength of relations with other families. These differences in overall correlational structure are illustrated in the relative boldness of the arrows in **Figure 1**.

The strong bidirectional connections between letter and phoneme nodes, as compared to those with semantic nodes, causes the letter-phoneme dynamic to cohere (resonate) before all others. These strong connections between letters and phonemes explain why phonology is so fundamental to reading and spelling. Stated differently, it explains why sound-alike words (ROSE and ROWS) may be confused in reading (Van Orden, 1987), and explains why the majority of spelling errors (ROZE instead of ROSE) are phonologically acceptable. (Van Orden & Goldinger, 1994, 1996 describe various other phenomena that derive from the powerful correlations of spelling and phonology.)

In a model analogous to **Figure 1**, presentation of a printed word activates letter nodes that, in turn, activate phoneme and semantic nodes. Following

initial activation, recurrent feedback begins among all these families of nodes. Similarly, presentation of a spoken word activates phoneme nodes that, in turn, activate semantic and letter nodes. (And, word production would begin with activation of semantic nodes that, in turn, activate phoneme and letter nodes.) In all these cases, initial activation leads to recurrent feedback among all families of nodes. However, the strongest recurrent dynamic is between letter and phoneme nodes, which creates the common basis of reading and spelling.

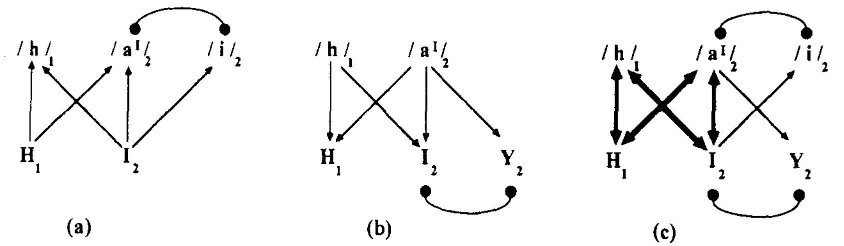

FIGURE 2. A simplified illustration of microdynamics that describe reading and spelling performance for the word HI. **a.** Presented with HI, activation feeds forward from letter nodes to phoneme nodes. **b.** In turn, phoneme nodes feed activation back to letter nodes. **c.** A resonance that emerges between letter and phoneme nodes corresponding to HI. (To reduce the number of lines in the figure, bidirectional connections are depicted with double-headed arrows.)

Figure 2 illustrates microdynamics. We zoom in on the connectivity between letter and phoneme nodes (and ignore, for now, phoneme-semantic and letter-semantic connectivity). In **Figure 2A**, reading the printed word HI activates the letter nodes H_1 and I_2, which activate the phoneme nodes $/h/_1$ and $/a^I/_2$ but also competing nodes such as $/I/_2$ (as in /hIt/) which must be inhibited. (The subscripts refer to the positions of the letters or phonemes within words.) **Figure 2B** shows how, in turn, phoneme nodes feed activation back to letter nodes (illustrated for the phoneme nodes $/h/_1$ and $/a^I/_2$). The phoneme node $/a^I/_2$ activates the correct letter nodes H_1 and I_2 and also competing letter nodes such as the letter node Y_2 in MY or BY. Thus, early patterns of activation are loosely structured, entailing activation of correct, but also many incorrect, candidates for resonance (Van Orden et al, 1990). Interactions between nodes then select combinations of nodes through cooperative-competitive dynamics (Grossberg & Stone, 1986; Stone & Van Orden, 1994).

Reliable performance emerges if the overall bidirectional configuration of connections favors mutual activation between the letter nodes H_1 and I_2 and the phoneme nodes $/h/_1$ and $/a^I/_2$. This advantage grows over time as the "strong grow stronger" and the "weak grow weaker" (McClelland & Rumelhart, 1981). This is illustrated in **Figure 2C**, which combines the flow of activation from letter nodes to phoneme nodes and from phoneme nodes back to letter nodes, as assumed in a recurrent network. Presentation of the spoken word $/ha^I/$ to the network (as in a spelling task) leads to a similar dynamic between phoneme and letter nodes. Thus, activation initiated in phoneme nodes may generate a coherent pattern of activity across letter nodes.

Why Spelling is More Difficult than Reading

In the course of writing, everybody experiences occasional doubts about how to spell a word, but we almost never forget how to read a word. This asymmetry between reading and spelling is evident at all levels of skill (Bosman & Van Orden, in press). The account we offer explains why people find spelling more difficult than reading. It may be simply described with respect to the previous illustrations of microdynamics (letter-phoneme dynamics), and macrodynamics (dynamics among families of nodes).

Returning to **Figure 2**, reading the word HI not only activates the phoneme nodes $/h/_1$ and $/a^I/_2$, and the letter nodes H_1 and I_2, but also all possible pronunciations of H_1 and I_2 and all possible spellings of $/h/_1$ and $/a^I/_2$. Thus, correctly reading a word requires inhibition of incorrect phoneme nodes, and correctly spelling a word requires inhibition of incorrect letter nodes. In the case of reading, the letters are presented to the model (or reader) such that phoneme→letter ambiguity is quite unlikely to produce full activation of incorrect letter nodes, because the persistent and stable input at letter nodes accelerates the formation of feedback loops with phoneme and semantic nodes (as illustrated by bold arrows in **Figure 2C**). However, in the case of spelling, this resonant pattern must be generated from phonologic and semantic activation alone. There is no environmental support for correct letter nodes.

The crux of spelling is that English orthography, generally, has more possible spellings for any given word than possible readings, and this is true for most (but not all) alphabetic writing systems (e.g., Stone, Vanhoy, & Van Orden, in press). Consider, for example, the multiple possible "spelling bodies" that may correspond to the "rime" /_ûrch/: _IRCH as in BIRCH, _ERCH as in PERCH, _URCH as in LURCH, and _EARCH as in SEARCH. Stone et al (in press) estimated that 69% of low-frequency English one-syllable words are spelling→phonology consistent (at the grain size of spelling bodies and rimes), but only 28% are phonology→spelling consistent (at the same grain size). In a larger sample, including both low- and high-frequency one-syllable words, 72% of all spelling→phonology consistent words were phonology→spelling inconsistent. These linguistic analyses clearly indicate that phonology→spelling inconsistency is the rule for English.

Although both reading and spelling are powerfully constrained by the strong correlational structure of the letter-phoneme dynamic, the occasional inconsistencies in these relations are resolved by different sources of constraint. Now,

we refer again to the illustration of macrodynamics in **Figure 1**. When a model "reads" a low-frequency, spelling→phonology *inconsistent word* such as PINT, the more consistent letter-phoneme relation would rhyme with MINT (and HINT, LINT, TINT, etc.; Kawamoto & Zemblidge, 1992). Similarly, the letter-phoneme dynamic would yield two correct pronunciations for words like WIND (although it would typically favor the more regular pronunciation, Kawamoto & Zemblidge, 1992). In both these cases, relatively strong semantic-phoneme relations may supply sufficient secondary constraints to encourage the appropriate letter-phoneme dynamic. In the case of WIND, semantic constraints may also be due to context, and contextual sources of semantic activation contribute via the relatively strong connections between semantic and phoneme nodes.

In the case of spelling, however, a model must resolve the inverted patterns of ambiguity in the phoneme-letter dynamic. To spell a low-frequency phonology→spelling inconsistent word such as HEAP, the rime /_ip/'s correct spelling would compete with a more strongly correlated incorrect spelling body _EEP (as in DEEP, BEEP, KEEP, PEEP, SEEP, and WEEP). Additionally, the phoneme-letter dynamic yields two correct spellings for homophones (e.g., ROSE/ROWS). In either case, correct spelling must rely on relatively weak semantic-letter dynamics to sufficiently activate the appropriate letter nodes (as illustrated in **Figure 1**); even contextual support is filtered through the weak letter-semantic connections. This weaker support for spelling, compared to the strong support for reading (i.e., phoneme-semantic dynamics) is the "macro-basis" for the asymmetry between reading and spelling. Thus, spelling is more difficult than reading for two reasons: (1) Microdynamic phoneme→letter relations are more inconsistent than letter→phoneme relations (Stone et al, in press), and (2) macrodynamic support for spelling (i.e., letter-semantic connections) is weaker than macrodynamic support for reading (i.e., phoneme-semantic connections).

Productive Use of This Simple Model

The recurrent dynamic system described above predicts a rather non-intuitive micro-effect. This effect concerns the consistency of relations between letters and phonemes. Until recently, all discussion of consistency has concerned a classic "feedforward," spelling→phonology effect. Inconsistent words such as PINT are named more slowly than consistent words such as DUCK. (_INT in PINT may be pronounced as in MINT; _UCK is only pronounced as in DUCK.) The feedforward consistency effect answers the question: Does it matter in word perception that a spelling may have more than one pronunciation? From most theoretical perspectives, this is the only sensible question. In a naming task, the letter string is unambiguous to subjects (it is right in front of their eyes); the only potential ambiguity arises with respect to derived phonology. However, our "feedback hypothesis" generalizes perceptual ambiguity in

the phonology→spelling direction as well. We were led to ask the feedback question: Does it matter in *visual* word perception that a *pronunciation* may have more than one *spelling*?

Recently, Stone et al (in press) tested for the effects of both feedforward and feedback consistency on performance in a lexical-decision task. They used a factorial design that included four types of words. In bidirectionally consistent words such as DUCK, the spelling body (_UCK) can only be pronounced one way, and the pronunciation rime (/_uk/) is only spelled one way. In spelling→phonology inconsistent words such as MOTH, the spelling body can be pronounced in multiple ways (BOTH), but the pronunciation rime (/_ôth/) is only spelled one way. In phonology→spelling inconsistent words such as HURL, the spelling body is pronounced in only one way, but the pronunciation rime can be spelled in more than one way (GIRL). In bidirectionally inconsistent words such as WORM, the spelling body can be pronounced in multiple ways (DORM), and the pronunciation rime can be spelled in multiple ways (FIRM). Stone et al found strong evidence for perception as a "two-way street;" correct response times were equally (and strongly) slowed by both feedforward and feedback inconsistency. Additionally, they found a reliable interaction; all inconsistent words produced approximately equal response times, even those that were inconsistent in both directions. Only words that were bidirectionally consistent produced better performance. Recently, Patrice Gibbs (personal communication, May 1995) found a similar effect of phonology→spelling consistency in a naming task.

Again, note the non-intuitive nature of this phenomenon. The letter string is clearly visible to the subject, and it remains visible until a response is recorded. However, if feedback from phonology suggests that some *other* letter string could have been presented, performance is slower. Ziegler and Jacobs (1995) reported a similar counter-intuitive finding in a letter-search task. Subjects in their experiment were briefly presented with a letter string such as BRANE (a "pseudohomophone" of the word "Brain"), followed by a pattern mask (#####). The subjects were instructed to indicate whether a predesignated letter, for example the letter "i," was present in the masked letter string. In the case of BRANE, they *mis*reported having seen the letter "i" more often than in a control stimulus. Similarly, they *mis*reported not having seen the letter "i" in the letter string TAIP (a pseudohomophone of the word "Tape"). Presumably, the phonology of the pseudohomophones BRANE or TAIP suggested that "Brain" or "Tape" was presented, causing subjects to misreport the presence or absence of the letter "i."

Developmental Dyslexia

Dyslexic children read poorly relative to non-dyslexic children of the same age, background, intelligence, and instructional level. No skill is more essential than reading for normal functioning within literate cultures; developmental

dyslexia can broadly undercut a child's potential for success and happiness (Bryant & Bradley, 1985). Dyslexics typically show qualitative differences from non-dyslexics in simple reading and language tasks. Moreover, their performance is impaired even relative to younger non-dyslexic children who successfully read at the same level (i.e., *reading-age* control subjects; Bosman, van Leerdam, & de Gelder, 1995; Pennington, Van Orden, Smith, Green, & Haith, 1990; see Rack, Snowling, & Olson, 1992, for review). These developmental dyslexics show specific deficits on tasks that require constructive use of phonology (e.g., phonological awareness and pseudoword-naming tasks, described shortly). We have offered an account of developmental dyslexia that derives from our account of phonologic mediation in skilled reading (Van Orden & Goldinger, in press).

A pseudoword-naming task requires fine-grain "phoneme-size" knowledge of how letter strings translate into phonology. In this task, a subject is shown a letter string, such as the pseudoword BINT, that shares spelling structure with actual words. (Consider MINT, BIN, etc. Skilled readers pronounce pseudowords analogously to words; Seidenberg, Plaut, Petersen, McClelland, McRae, 1994). Dyslexic readers name pseudowords much more slowly and produce more "unacceptable" pronunciations than reading-age controls (Rack et al, 1992). They may also perform poorly when judging whether *someone else* has given an "acceptable" pronunciation of a pseudoword (Snowling, 1980). Poor performance in pseudoword naming is a primary symptom of dyslexia.

Correct performance in phonological-awareness tasks also depends upon fine-grain, phoneme-size knowledge of the phonology of a word. These tasks typically require subjects to manipulate or judge the phonology of words. In a "pig Latin" task, for example, the first phoneme of a word must be moved to the end and pronounced with /AY/ (e.g., /dog/ becomes /og·day/). Dyslexics perform very poorly on this task relative to reading-age control subjects, even when they need only recognize whether someone else has produced correct pig Latin (Pennington et al, 1990). Deficits in phonological awareness are typically correlated with deficits in pseudoword naming, and both deficits appear to be influenced by heredity (Olson, Wise, Conners, Rack, & Fulker, 1989; Pennington et al, 1990).

These findings all motivate the hypothesis that dyslexia is a deficit in fine-grain knowledge of phonology and its relation to print in alphabetic languages. The importance of phonology in dyslexia agrees with our account of skilled reading in which phonology also plays a crucial role. The crux of reading is perception of individual printed words (Perfetti, 1985), and the crux of word perception is coherent phonology. Accordingly, developmental dyslexia might be explained by an absence of phonology in reading. It turns out, however, that *absent* phonology is far too simple a hypothesis (Bruck, 1988). Dyslexic

subjects, who show a pronounced deficit in pig-Latin performance (Pennington et al, 1990), nevertheless produce a very high proportion of categorization errors to homophonic foils (e.g., ROWS or ROZE categorized as FLOWERS, Van Orden et al, 1990; Van Orden & Goldinger, 1996).

The paradox for the absent-phonology hypothesis is that the same dyslexic subjects show *both* negative and positive phonology effects. Demonstrations of negative phonology (such as pseudoword naming and pig-Latin deficits) are consistent with the absent-phonology hypothesis (Bruck, 1988). However, categorization errors to homophonic foils are not. A key difference between pseudoword (BINT) naming versus categorization may be the added constraints in categorization produced by category names. This source of constraint may exaggerate the dyslexics' susceptibility to phonology in the categorization task, especially with pseudoword homophones such as ROZE. It does so, however, by enhancing ("cleaning up") letter-phoneme dynamics, which explains their very high error rates to homophonic foils. Thus, our ability to observe phonology effects in dyslexics is partly a function of the task examined. We propose that different tasks emphasize different grain sizes of phonology, and these contribute to the respective positive and negative effects. To better understand our proposal, it is first necessary to understand how *covariant learning* serves as a basis for non-dyslexic reading.

Crosstalk is the basis of covariant learning. Crosstalk extracts positive correlations between families of nodes (cf. Reeke & Edelman, 1984). Reading performance is enhanced by consistent crosstalk whenever a letter-phoneme or letter-phoneme-semantic correspondence is shared across a neighborhood of words. In word naming, consistent crosstalk is the source of many common effects, such as rule-strength and word-frequency effects. Rule strength is estimated by a count of all words that share a particular letter-phoneme correspondence. Strong-rule words are composed of letter-phoneme relations that appear in many words (K-/k/). Weak-rule words have at least one letter-phoneme relation that is relatively rare (ZZ-/z/). Strong-rule words (DESK) and pseudowords (DASK) are named faster and more accurately than weak-rule words (FIZZ) and pseudowords (NOZZ; Rosson, 1985). Also, high-frequency words are named faster and more accurately than low-frequency words (Forster & Chambers, 1973). **Figure 3** illustrates how these effects would emerge via covariant learning in a very simple model. BE and BY share a relatively strong rule (B-/b/), and BE is the more frequent word (in the figure).

In **Figure 3A**, a BE learning trial brings four pairs of nodes into collective resonance: $B_1 \longleftrightarrow /b/_1$, $B_1 \longleftrightarrow /i/_2$, $E_2 \longleftrightarrow /b/_1$, and $E_2 \longleftrightarrow /i/_2$. Such resonance increases the connection weights between all the nodes involved. At this point in the model's development, the resonance $B_1E_2 \longleftrightarrow /b_1i_2/$ is an "encapsulated" whole. Although we can anticipate potential subresonances in the *a*

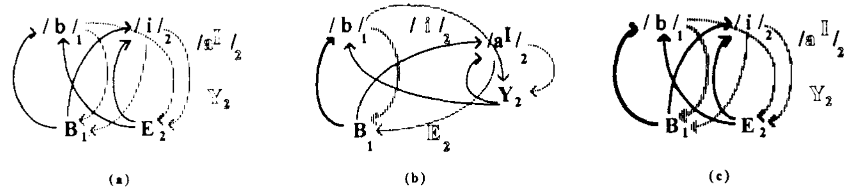

FIGURE 3. Consistent crosstalk in a recurrent network capable of covariant learning. **a.** The consequences for the connections between letter nodes and phoneme nodes of a learning trial for the word BE. **b.** Likewise for the word BY. **c.** A second learning trial for the word BE. The presence of a line between two nodes indicates an increase in the connection weights between them. The width of the lines ranks the strength of the relations that accumulates across learning trials. Notice across the figures that the width of the lines between B_1 and $/b/_1$ increases faster than those of other relations. This is due to consistent crosstalk.

priori relations between letter and phoneme nodes, these are not reflected in the model's behavior. If the network were presented again with the word BE, the activation values of the same four nodes would grow symmetrically toward resonance. Thus, the model's behavior would reflect only coarse-grain, word-size knowledge. This simple figure illustrates how dimensionally nonspecific (holistic) relations can emerge behaviorally prior to relational rule-like knowledge (Thelen & Smith, 1994). The later emergence of rule-like knowledge is illustrated in panels **3B** and **3C**.

A subsequent BY trial, shown in **3B**, brings its four pairs of nodes into collective resonance ($B_1 \longleftrightarrow /b/_1$, $B_1 \longleftrightarrow /a^I/_2$, $Y_2 \longleftrightarrow /b/_1$, and $Y_2 \longleftrightarrow /a^I/_2$), thus adjusting all the connection weights involved. Notice at this point that the connections between B_1 and $/b/_1$ have been adjusted more often than any other connections. If, in turn, another BE learning trial occurs (**3C**), then the four pairs of connection weights: $B_1 \longleftrightarrow /b/_1$, $B_1 \longleftrightarrow /i/_2$, $E_2 \longleftrightarrow /b/_1$, and $E_2 \longleftrightarrow /i/_2$ are adjusted again by its four component resonances. Because B correlates with the same pronunciation in BY and BE, the configuration of weights in the resonance $B_1 \longleftrightarrow /b/_1$ is tuned toward this strong subword resonance more often than configurations promoting other component resonances. The bidirectional connections between B_1 and $/b/_1$ emerge as a strong rule via consistent crosstalk.

Strong-rule resonances, such as $B \longleftrightarrow /b/$, are examples of local (fine-grain) dynamics exhibiting relatively high *self-consistency*. After learning, the component resonances of a strong-rule word show themselves behaviorally because they coalesce quickly and thereby facilitate naming. Consequently, even relatively unfamiliar words are named quickly if they are composed of

strong rules (Rosson, 1985). Additionally, the naming of pseudowords (e.g., BINT) is primarily constrained by these same fine-grain resonances, as will be seen in our account of developmental dyslexia.

Earlier in this chapter, we described a fictitious nervous system to introduce the construct *resonance*. We also oversimplify the nervous system in this section to introduce our account of dyslexia. Postmortem studies have found anatomical anomalies in the brains of dyslexics (e.g., see Galaburda, Rosen, & Sherman, 1989) that may be due to subtle anomalies in neuronal migration. Small deviations in neural positioning may cause large changes in patterns of connectivity between neurons in different brain regions. A rough analogy with connectivity in network models inspired our "haphazard-connections" hypothesis concerning the performance deficits of dyslexics. (Please do not interpret this rough analogy as a claim to anatomical plausibility. We merely wish to acknowledge the inspiration for our behavioral account.)

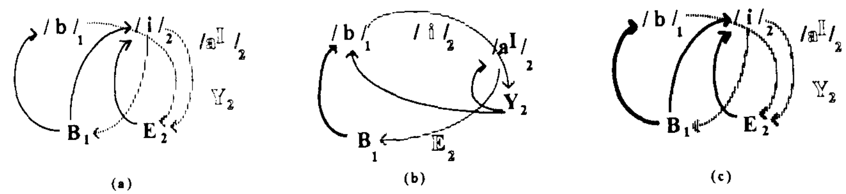

FIGURE 4. An illustration of the failure to develop fine-grain dimensions of word perception. **a.** The consequences for the connections between the letter nodes and phoneme nodes of a learning trial for the word BE when letter and phoneme nodes are haphazardly connected. **b.** Likewise for the word BY. **c.** A second learning trial for the word BE. The presence of a line between two nodes indicates an increase in the connection weights between them. Notice that the bidirectional fine-grain relation between B_1 and $/b/_1$ does not emerge in this haphazardly connected illustration.

Figure 4 illustrates a haphazard pattern of connectivity between letter and phoneme nodes. (Compare this to **Figure 3**, in which all connections are symmetrical.) We can easily track the outcome of covariant learning, given this haphazard connectivity. Once again, consistent crosstalk extracts positive correlations between letter and phoneme nodes. However, although BE and BY share a relatively strong rule (B-/b/), it does not emerge as a self-consistent subresonance in the behavior of the model. This is the key to our account.

In **Figure 4A**, a BE learning trial brings four nodes into collective resonance ($B_1 \rightarrow /b/_1$, $B_1 \longleftrightarrow /i/_2$, $E_2 \leftarrow /b/_1$, and $E_2 \longleftrightarrow /i/_2$), and adjusts six connection weights. As in **Figure 3A**, at this point in the model's development the resonance $B_1E_2 \longleftrightarrow /b_1i_2/$ is an encapsulated whole. The model would

correctly produce the whole-word phonology of BE, but its behavior would be opaque to substructures such as $B_1 \longleftrightarrow /b/_1$. A subsequent BY trial, shown in **4B**, brings four pairs of nodes into collective resonance ($B_1 \longleftrightarrow /b/_1$, $B_1 \leftarrow /a^I/_2$, $Y_2 \longleftrightarrow /b/_1$, and $Y_2 \rightarrow /a^I/_2$) and adjusts five connection weights. Notice the difference between panel **B** in **Figure 3** and panel **B** in **Figure 4**. The feedforward connection weight $b_1 \rightarrow /b/_1$ grows faster than other connections in the resonance $B_1E_2 \longleftrightarrow /b_1i_2/$, due to covariant learning. However, the subword relation $B_1 \longleftrightarrow /b/_1$ does not grow in self-consistency due to the absence of feedback connections. The haphazard connectivity between B_1 and $/b/_1$ in **Figure 4B** does not allow the emergence of the component resonance $B_1 \longleftrightarrow /b/_1$, and this state of affairs is unchanged by additional BE trials (**4C**).

Consider the subword dynamics for the word BE following the learning trials depicted in **Figure 4**. An advantage due to covariant learning between B_1 and $/b/_1$ may still affect performance. The relatively strong connection $B_1 \rightarrow /b/_1$ would promote a faster overall time to resonance for the word BE. The node $/b/_1$ is strongly activated by B_1, and $/b/_1$ conserves this strong activation for the whole-word resonance when it feeds activation back to E_2. In turn, E_2 feeds this activation forward to $/i/_2$, $/i/_2$ feeds it back to B_1 and E_2, and the activation has been conserved in a *word-size feedback loop*. At no time, however, does a *subresonance* $B_1 \longleftrightarrow /b/_1$ emerge in the behavior of the model.

We noted previously that pseudoword (BINT) naming is primarily dependent upon fine-grain resonances such as $B_1 \longleftrightarrow /b/_1$. These are necessary to insure integrity of pseudoword pronunciation, but they do not reliably emerge in a model with haphazard connectivity. This translates behaviorally into a deficit in pseudoword naming (as well as reading proper names, unfamiliar words, etc.). Skilled pseudoword naming is a fairly predictable function of the statistical relation between words' spellings and their pronunciations (Seidenberg et al, 1994). With haphazard connectivity, a model is robbed of its strongest source of information about how pseudoword spellings relate to phonology. A model analogous to our simple illustration would fail to derive a full complement of the fine-grain letter-phoneme relations necessary to mimic skilled pseudoword naming. Thus, dyslexia may be a failure of perceptual development to derive these fine-grain dimensions of word perception.

Remember, however, that strong positive phonology effects are found in categorization performance with both word (ROWS) and pseudoword (ROZE) homophones. The word effect is explained by the coarse-grain (word-size) phonology that emerges in the haphazardly connected model. However, the effect of pseudoword (ROZE) phonology requires the added contextual constraints on phoneme-semantic dynamics due to availability of the category name (FLOWER). This context is strong enough to cause skilled readers to misinterpret highly familiar homophone words in a categorization task (Jared & Seidenberg, 1991). We suggest that context acts to compensate for the noisy letter-phoneme dynamics of dyslexic readers via semantic-phoneme dynamics, as illustrated in **Figure 1**.

The essential point for implementing our scheme is that haphazard connectivity blocks the emergence of letter-phoneme resonances (i.e., attractors; see Jordan, this volume) that would normally act relatively independently of the coarse-grain resonances from which they derive. These fine-grain dynamic structures are necessary to mimic the full range of intact performance to printed language. We could even zoom in on the phoneme nodes, revealing their finer-grain resonances between acoustic and articulatory nodes. Then, we could propose haphazard connectivity between acoustic and articulatory nodes, thus precluding the development of proper "phoneme resonances." In this way, the haphazard-connectivity hypothesis might be extended to explain poor performance on phonological-awareness tasks.

Acquired Dyslexia

Patients with *acquired dyslexia* have reading difficulties as a consequence of brain trauma. In a seminal article, Marshall and Newcombe (1973) described two apparently distinct syndromes of acquired dyslexia: *surface* and *deep* dyslexia. Both are defined by characteristic profiles of errors in the naming task (Shallice, 1988). The utility and reliability of syndrome categories is highly controversial (Caramazza, 1986), but we need not endorse such distinctions for our purpose. We merely focus on theoretically important patterns of naming errors associated with each syndrome. Our goal is to produce similar errors in "lesioned" models that previously produced skilled patterns of naming performance. Here, we describe briefly our simulations of two error types identified by Marshall and Newcombe: the regularization error and the semantic error.

Regularization errors are characteristic of surface-dyslexic patients. These errors occur when words such as PINT, with irregular pronunciations, are incorrectly read aloud to rhyme with similar regular words (e.g., MINT, HINT, and LINT). Although skilled readers also occasionally make regularization errors (Kawamoto & Zemblidge, 1992), surface-dyslexic patients make many more. *Semantic errors* are characteristic of deep-dyslexic patients, occurring when words are incorrectly read aloud as semantically related words. For example, the word BUSH might be read aloud incorrectly as TREE. The separate occurrence of semantic and regularization errors is sometimes interpreted as evidence against recurrent-network models (Shallice, 1988; see Van Orden et al, 1996 for a review and counter-argument). Farrar and Van Orden (1994) recently simulated these two error types.

We began with a recurrent-network architecture very similar in structure to the simple illustrations presented earlier in this chapter. Three families of nodes (see **Figure 1**) were "taught" a sample of English words using a Hebbi-

an-type learning algorithm, until the model produced patterns of naming performance similar to those of skilled readers. (Specifically, we implemented 10 learning trials for each "high-frequency" word and one for each "low-frequency" word. The "naming response" was taken from the pattern of the most active phoneme units, and "naming time" was defined as the number of cycles required to generate a coherent pattern of activation of the phoneme units.) Of particular present relevance, the model produced an interaction of frequency with consistency. Low-frequency inconsistent words such as PINT were named more slowly than low-frequency consistent words such as DUCK, whereas all high-frequency words were named quickly (see Waters & Seidenberg, 1985).

To simulate the regularization error, we added noise to our intact network. Noise was implemented as a uniform distribution of small positive or negative changes in activation that were added in each cycle to the activation values of randomly chosen nodes. The effect was to erode the network's capacity for correct naming of words having highly inconsistent pronunciations, such as PINT. Instead of the correct phonemes for PINT, the network activated phoneme nodes that regularized PINT to rhyme with MINT. In effect, the noise eroded the (already weak) phoneme-semantic constraints, such that the model expressed only the powerful constraints of letter-phoneme dynamics. Because letter-phoneme dynamics primarily reflect the strongest correlations between letters and phonemes, these dynamics lead to regularization errors (Kawamoto & Zemblidge, 1992).

It is important to note that we could have implemented noise in many ways with the same consequences. For example, we could have introduced small changes in randomly chosen connection strengths. Similarly, the *locus* of noise is not crucial. Bidirectional flow of activation makes the system highly interdependent. Consequently, noise introduced anywhere in the system spreads throughout the system, in the next time step.

To simulate semantic errors, we further "lesioned" the noisy network that was producing regularization errors. We set all of the letter-phoneme connection weights to zero, effectively "cutting" those connections. (We could have cut fewer connections with the same effect; the minimum proportion of disconnections that produces semantic errors is interdependent with other modeling choices.) Subsequently, the network produced semantic errors; when presented with BUSH the network generated a pattern of activity across phoneme nodes corresponding to TREE. Setting the letter-phoneme connections to zero creates a highly unstable network, causing it to rely on semantic-phoneme dynamics, the most reliable remaining source of constraints. However, in the absence of letter-phoneme constraints, semantic-phoneme dynamics are sometimes misled into a semantic error from which the normally weak letter-semantic dynamics cannot rescue the network. Semantic errors are especially likely when semantic

nodes of one word (BUSH) are strongly correlated with phoneme nodes of a different word (TREE).

Mind as Embodied Action

The attentive reader has probably noticed that we refrained from calling our cognitive account a neural network. We chose the term recurrent network instead. This choice was not made simply for aesthetic reasons. The term *neural* in "neural network" has a connotation we wish to avoid. It suggests too strong an analogy with the nervous system, or (worse yet) that cognition should be explained in terms of the nervous system. For all we know, there may be more differences than similarities between cognitive behavior and the behavior of nervous systems (cf. Freeman, 1995). Thus, "nodes" in our recurrent network do not refer to neurons, nor do their interconnections refer to synapses.

The previous caveat resurrects the issue of the metaphor underlying our account, specifically, *mind as embodied action* (see Introduction). This view of cognition, as described by Varela et al (1991), means, first, that

> ...cognition depends upon the kinds of experience that come from having a body with various sensorimotor capacities, and second, that these individual sensorimotor capacities are themselves embedded in a more encompassing biological, psychological, and cultural context. (pp. 172-173)

The term *action* emphasizes that perception and action are fundamentally interrelated (see also Turvey & Carello, 1981). This proposed interdependence is obvious in the behavior of a recurrent network. When our network is presented with a printed word, the initial activations of the letter nodes feed activation forward to phonologic and semantic nodes that, in turn, feed activation back to the letter nodes. In a strongly nonlinear system, feedforward and feedback activation undergo successive nonlinear transformations, eventually producing a resonant whole. This illustrates, in model systems, the irreducible interdependence of input and output (see also van Leeuwen, Steyvers, & Nooter, 1995).

It is easy to confuse nodes in a network (or *subsymbols*, Van Orden et al 1990; Van Orden & Goldinger, 1994) with traditional symbolic representations. This is another confusion that we wish to avoid. Nodes are not mental representations. They are pragmatic notations for purposes of modeling or illustration, and serve a narrative function only. Thus, nodes are not psychologically real (atomic) units of cognition, and they should not be attributed causal or explanatory properties independent of the dynamics in which they participate (cf. Putnam, 1981; Turvey & Carello, 1981). The network models that we propose refer only to observed patterns of intercorrelation between laboratory manipulations and performance. This pragmatic approach to model-

ing implies that nodes chosen at one grain size are no more "real" than nodes that might have been chosen at other grain sizes. To rephrase, modeling a particular behavior or performance requires a smart (or pragmatic) choice of nodes (see also Putnam, 1981). We choose nodes at the finest grain size of reliable covariation between laboratory manipulations and performance (Van Orden et al, 1990; Van Orden & Goldinger, 1994). A fairly good description of results obtained from reading and spelling research can be achieved using letter, phoneme, and semantic (feature) nodes. Note that a degree of arbitrariness is unavoidable; other research problems may lead to other choices. The crux is the ability to mimic, in as parsimonious a model as possible, the observed complexity in laboratory performance.

The pragmatic constraints on the choice of nodes are further revealed by the following example. Choosing to model letter perception using a grain size of letter nodes would be too "coarse" a choice, because it ignores reliable effects of font, handwriting, and other episodic variables (e.g., Sanocki, 1987). A finer grain size (e.g., letter-feature nodes) would be necessary. Conversely, choosing to model sentence comprehension using a grain size of letter nodes is a too "fine" a choice of grain. Using letter nodes to model the phenomena of sentence comprehension would build in unnecessary detail. These performance phenomena are typically word- or morpheme-size, and nodes chosen at these grain sizes would be more appropriate. The strength of our model comes not from the discovery of true nodes, but from the generality and simplicity of its behavioral account. Dynamic interactions among small families of nodes can account for a vast literature of performance in laboratory reading tasks, and the entailed principles extend to cognition at large (Stone & Van Orden, 1994; Van Orden et al, 1990; Van Orden & Goldinger, 1994).

In summary, we have shown that a simple recurrent network has great utility for mimicking phenomena in reading and spelling. Our approach motivates principled explanations for why phonology is fundamental to reading and spelling, why spelling is more difficult than reading, why words with multiple rime spellings are more slowly read, and why developmental and acquired dyslexics have difficulties in reading.

ACKNOWLEDGMENTS

Preparation of this chapter was funded by research grants 5 R29 NS26247-05 from the National Institute of Neurological Disorders and Stroke, National Institutes of Health, awarded to Guy Van Orden; 1 R29 DC02629-01 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health, awarded to Steve Goldinger; 1 F32 DC 00187-01 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health, awarded to Bill Farrar, and a NATO-Fellowship N58-92 from the Netherlands Organization for Scientific Research (NWO) awarded to Anny Bosman. Correspondence should be addressed to the senior author.

REFERENCES

Note: The chapter numbers in which each reference is cited are indicated in parentheses following the reference.

Abbott, L.F., Varela, J.A., Sen, K., & Nelson, S.B. (1997). Synaptic depression and cortical gain control. *Science*, **275**, 220-224. (Pt 5)

Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. Cambridge, UK: Cambridge University Press. (8)

Abeles, M., Bergman, H., Margalit, E., & Vaadia, E. (1993). Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. of Neurophysiology*, **70**, 1629-1638. (8)

Abraham, W.C., Bliss, T.V.P., & Goddard, G.V. (1985). Heterosynaptic changes accompany long-term but not short-term potentiation of the perforant path in the anaesthetized rat. *J. of Physiology*, **363**, 335-349. (6)

Abraham, W.C., Dragunow, M., & Tate, W.P. (1991). The role of immediate early genes in the stabilization of long-term potentiation. *Molecular Neurobiology*, **5**, 297-314. (5)

Abraham, W.C., Mason, S.E., Demmer, J., Williams, J.M., Richardson, C.L., Tate, W.P., Lawlor, P.A., & Dragunow, M. (1993). Correlations between immediate early gene induction and the persistence of long-term potentiation. *Neuroscience*, **56**, 717-727. (6)

Ackerman, R., Finch, D., Babb, T., & Engel, J. (1984). Increased glucose metabolism during long-duration recurrent inhibition of hippocampal pyramidal cells. *J. of Neuroscience*, **4**, 251-264. (24)

Ackley, D., & Littman, M. (1992). Interactions between learning and evolution. In C.G. Langton, C. Taylor, J.D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II: Proceedings of the Workshop on Artificial Life* (pp. 487-509). Redwood City: Addison-Wesley. (4)

Agnati, L.F., Zoli, M., Pich, E.M., Benfenati, F., & Fuxe, K. (1990). Aspects of neural plasticity in the central nervous system VII: Theoretical aspects of brain communication and computation. *Neurochemistry International*, **16**, 479-500. (5)

Agranoff, B.W., & Klinger, P.D. (1964). Puromycin effect on memory formation in the goldfish. *Science*, **146**, 952-953. (6)

Ahissar, E., Vaadia, E., Ahissar, M., Bergman, H., Arieli, A., & Abeles, M. (1992). Dependence of cortical plasticity on correlated activity of single neurons and on behavioral context. *Science*, **257**, 1412-1415. (8)

Aiple, F., & Krüger, J. (1988). Neuronal synchrony in monkey striate cortex: Interocular signal flow and dependency on spike rates. *Experimental Brain Research*, **72**, 141-149. (8)

Akshoomoff, N.A., Courchesne, E., Pres, G.A., & Iragui, V. (1992). Contribution of the cerebellum to neuropsychological functioning: Evidence from a case of cerebellar degenerative disorder. *Neuropsychologia*, **30**, 315-328. (24)

Albright, T.D., & Gross, C.G. (1990). Do inferior temporal cortex neurons encode shape by acting as Fourier descriptor filters? *Proceedings of International Conference on Fuzzy Logic and Neural Networks*, Izuka, Japan, 375-378. (9)

Albus, J. (1981). *Brains, behavior, and robotics*. Peterborough, NH: Byte Books. (15)

Alexander, G.E., Crutcher, M.D., & DeLong, M.R. (1990). Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. In H.B.M. Uylings, C.G. VanEden, J.P.C. De Bruin, M.A. Corner, & M.G.P. Feenstra (Eds.), *The prefrontal cortex: Its structure, function and pathology* (pp. 119-146). Amsterdam: Elsevier.

Alexander, G.E., DeLong, M.R., & Strick, P.L. (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, **9**, 357-381. (17)

Alkon, D.L., Amaral, D.G., Bear, M.F., Black, J., Carew, T.J., Cohen, N.J., Disterhoft, J.F., Eichenbaum, H., Golski, S., Gorman, L.K., Lynch, G., McNaughton, B.L., Mishkin, M., Moyer, J.R.J., Olds, J.L., Olton, D.S., Otto, T., Squire, L.R., Staubli, U., Thompson, L.T., & Wible, C. (1991). Learning and memory. *Brain Research Review*, **16**, 193-220. (6)

Allen, R.B. (1989). Adaptive training of connectionist state machines. Paper presented at ACM Computer Science Conference, Louisville, KY. (16)

Allman, J. (1990). Evolution of the cortex. In A. Peters & E.G. Jones (Eds.), *Cerebral cortex, Vol. 8a: Comparative structure and evolution of cerebral cortex, Part I* (pp. 269-283). New York: Plenum Press. (4)

Altman, J. (1969). Autoradiographic and histological studies of postnatal neurogenesis IV: Cell proliferation and migration in the anterior forebrain, with special reference to persisting neurogenesis in the olfactory bulb. *J. of Comparative Neurology*, **137**, 433-458. (2)

Altman, J., & Bayer, S.A. (1990). Mosaic organization of the hippocampal neuroepithelium and the multiple germinal sources of dentate granule cells. *J. of Comparative Neurology*, **301**, 325-342. (3)

Altman, J., & Bayer, S.A. (1995). *Atlas of prenatal rat brain development*. Boca Raton, FL: CRC Press. (3)

Alvarado, M., & Rudy, J. (1995). A comparison of kainic acid plus colchicine and ibotenic acid-induced hippocampal formation damage on four configural tasks in rats. *Behavioral Neuroscience*, **109**, 1052-1062. (12)

Amaral, D.G. (1987). Memory: Anatomical organization of candidate brain regions. N. Plum (Ed.,), *Handbook of physiology: Section I. Neurophysiology, Vol. 5: Higher functions of the brain* (pp. 211-294). Bethesda, MD: American Physiological Society. (18) (26)

Amaral, D.G., & Witter, M.P. (1989). The three-dimensional organization of the hippocampal