Computational approaches to normal and impaired spelling

Gordon D.A. Brown

Department of Psychology

University of Wales

Bangor

Gwynedd

LL57 2DG

and

Richard P.W. Loosemore

Department of Psychology

University of Warwick

Coventry

CV4 4AL

Brown, G.D.A. & Loosemore, R.P.W. (1995). Computational approaches to normal and impaired spelling. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and application.* Chichester, U.K.: John Wiley.

Introduction

The three chapters in this section of the book all describe "connectionist" or "neural network" models of spelling. In this chapter we give a brief introduction to connectionist modelling, and argue that adopting a computational approach can help us to develop better models of spelling and gain a different kind of understanding of the processes involved in normal spelling and the kinds of problems that could lead to impaired spelling development.

The rest of the chapter is devoted to a description of our own connectionist model of normal and dyslexic spelling development. We use the connectionist approach to investigate the difficulty, in computational terms, of learning to spell different types of words; types that differ in their sound-to-spelling characteristics. This leads to a new classification of word types, and to the realisation that it is insufficient simply to classify words as "regular" or "irregular" in their sound-to-spelling correspondence. We show that a simple connectionist model can learn to spell both regular and irregular words (as traditionally defined) with a reasonable degree of accuracy. We then describe the results of some experiments we have conducted which show that children, when asked to spell the same words that the model has learned to spell, exhibit the same type of difficulty on the same word type as does the model. In particular, words which contain sound-to-spelling correspondences which are either *unusual* (i.e., low frequency) or *irregular*, (i.e., with exceptions) cause particular difficulty both for children at various stages of development and for the model.

The fact that the level of difficulty experienced by the model on different word types so closely mirrors the pattern of difficulty found in children is taken as evidence that the process of learning to spell can usefully be viewed as one of mastering a set of statistical associations between representations of the phonological forms of words and representations of their orthographies. This kind of approach stands in contrast to "cognitive" or "information processing" approaches which refer to the "rules" or "strategies" that children use at different stages in learning to spell.

We also use the model to examine the question of whether developmentally dyslexic spelling problems can be characterised in terms of the amount of computational resource that dyslexics bring to the task of learning to spell. We show that reducing the computational capacity of the spelling learning network leads to slower and less accurate learning of both regular and irregular words. We briefly describe the results of an experiment on dyslexic spellers which demonstrates that the pattern of difficulty experienced by "dyslexic" versions of the model is very similar to that experienced by dyslexic children. In the final section, we use our own model and those reported in the subsequent two chapters (by Houghton et al. and by Olson and Caramazza) to draw out some general conclusions about the utility of computational modelling in general, and connectionist modelling in particular, to our understanding of normal and disturbed spelling development.

The need for computational models

Many of the chapters in this book refer to some version of the "dual-route" model of spelling, in which there are two separate routines available by which the correct spelling of a word maybe derived. Such models are generally verbally described, often with an accompanying diagram (see, e.g., chapter by Barry in the present volume). This kind of approach has proved useful in improving our understanding of both normal and impaired spelling development, as well as the acquired disorders of spelling sometimes observed in patients with acquired brain injuries. However, models may be more rigorous if they are expressed *formally* in some way. A formal specification of a model will usually mean either a mathematical description of the model and the way it works, or the implementation of the model in the form of a computer program.

In the present section we focus on computational implementation of models of spelling. There are several advantages to expressing models in the form of computer programs. First of all, such expression can provide a guarantee that the relevant model is workable, i.e., that it includes all the knowledge and mechanisms necessary for the system to perform the task that it is designed for. It is easy to be mislead by an elegant verbal description of a psychological model into the belief that the model is *explanatory* in that it has specified the relevant causal mechanisms, can exhibit the relevant behaviour and also give rise to novel predictions. However (as we ourselves have found to our cost in a number of areas of psychology), when one tries to write a computer program to implement even an apparently simple psychological model, it is common to find that many crucial details as to how the model is supposed to work are left unspecified. Thus if a verbally specified model is incomplete, it can be very difficult to detect the incompleteness without actually going through the discipline of attempting to write a computer program to express the model.

A second, and related, reason to implement models computationally concerns the derivation of predictions from the model. Implementing a model in the form of a computer program provides an effective guarantee that the behaviour of the model can be examined under any simulated conditions. Thus predictions can be systematically derived from the model. These can then be tested and possibly disconfirmed experimentally, thus ensuring that the model has true explanatory content rather than being nothing more than a redescription of data that already exists. When a model is incompletely specified, as is frequently the case with purely verbally described models, in contrast it is not always possible to determine unambiguously what the predictions of the model in particular circumstances will be. This can lead to problems in deciding whether or not to accept the results of a specific experiment as evidence for or against the model concerned.

Thirdly, a related consideration concerns the explanatory constructs that are used in the model. This applies to cognitive models in particular. For example, information processing models frequently refer to "strategies," "mental stages," "rules," "representations" and so on. Unfortunately it may not always be clear exactly what these constructs are, how they are supposed to work, and whether different researchers mean the same thing when they talk about them. Because these psychological constructs are internal they cannot be observed directly. Many people have pointed to the dangers in postulating internal mechanisms which have arbitrary complexity in an attempt to explain some aspect of behaviour, for there is a danger that the unobservable mechanisms that are proposed may not be explanatory at all and indeed may not be anything more than a redescriptions of the relevant data. If the relevant rules and processes form part of a computer program however, then there can be no disputes about what is meant by them, because it is possible for anyone to examine or recreate the program and see exactly how the postulating mechanisms are supposed to work.

A further reason for implementing models in the form of computer programs confirms the complexity and predictability of the relevant underlying processes. In the case of weather forecasting or the modelling of traffic flows, for example, the systems under study are so complex that it is impossible to predict their behaviour without simulating the system in some way. We believe it is likely that this will also turn out to be the case in psychology for models of spelling.

A further advantage of computational modelling is the increased insight given to the modeller as the nature of the problems occur. Parts of some process which intuitively seemed extremely easy to solve may turn out to be unexpectedly problematic when one is required to have a sub-routine in one's program to cope with them. Thus, for example, the true complexity of tasks like visual object recognition or auditory speech perception, which seem effortless to us, has been dramatically underlined by the failure of decades of attempts to build computer programs to carry out the same tasks as fast and accurately as humans. In many cases it is only by trying to imitate how the brain carries out a task that we have come to realise just how hard the task is to solve, and we suspect that the same may turn out to be true for models of spelling.

The three chapters in this section each demonstrate new insights as a direct result of the attempt to build computational models of spelling, and we do not discuss their conclusions further here. First we return to a brief description of the connectionist approach to the computational modelling of psychological processes.

Connectionism

Over the past decade or two there has been a great interest in building computational models of psychological processes that are based, more or less loosely, on the known structure of the brain. In computational terms, the brain can be viewed as being made up of a vast number of processing components (neurons or synapses) which are, in comparison with the central processing unit of modern digital computers, both relatively simple and relatively slow. Thus the computational power of the brain appears to come from its possession of a vast number of computing elements which are heavily inter-connected (each unit being connected to around 10,000 other ones on the average in the brain). Attempts to simulate brain-like architectures by building interconnected networks of simple artificial neurons (neural networks), have shown that such networks exhibit many psychologically plausible characteristics such as learning capability, the ability to generalise and respond to novel input, graceful degradation of performance under damage, flexible pattern perception and completion, and many others (see, e.g., papers in Rumelhart & McClelland, 1986a, and in McClelland & Rumelhart, 1986). The units in such models are generally simpler by far than real neurons in the brain, but the use of such models has nevertheless lead to greater understanding of many psychological processes.

How does a simulated neural network operate? Artificial neurons in connectionist models usually have an "activation value" associated with them, often between 0 and 1. This is sometimes though of as being analogous to the firing rate of a real neuron. Sometimes binary-valued artificial neurons are used, and these must take on the value of either 0 or 1. Psychologically meaningful objects (such as the pronunciation of words) can then be represented as patterns of this activity, i.e., the presence of 0's and 1's across sets of the artificial neurons. For example, the word SOAP might be represented by the pattern 101100011, while the word PILL was represented as the pattern 001011100. In our model of spelling development, to be described below, one sub-population of the units in the network is used to represent the pronunciations of words, and another sub-population is used to represent the orthographic forms of words.

The units in the artificial network are typically multiply interconnected by connections with variable strengths or *weights*. These connections permit the level of activity in any one unit to influence the level of activity in all the units that it is connected to. These connection strengths can be adjusted by a suitable learning algorithm, in such a way that when a particular pattern of activation appears across one population it can lead to a desired pattern of activity arising on another set of units. Thus, for example, when a pattern of activity is imposed on the subset of units in the spelling model that are used to represent the pronunciation of words, this can lead to the formation within the artificial network of a pattern of activity on the separate, but connected, sub-population of units that are designated to represent the spelled form of words. If the connection strengths have been set appropriately by the learning rule, then it may be possible for units representing the pronunciation of a particular word to cause the units that represent the correct spelling of that word to become activated. This means that the network could be said to have learned how to spell a word with that pronunciation.

Finally, these networks have the capacity to store many different associations within the same artificial network. These properties of the networks have been extensively described and explored elsewhere (e.g., Rumelhart & McClelland, 1986a; McClelland & Rumelhart, 1986).

We now attempt to explain the principles behind our own connectionist model of spelling, which in general terms works the same way as the generic connectionist network described above. The architecture of the model, and the form of representations it uses, are in many respects similar to the connectionist model of reading developed by Seidenberg and McClelland (1989a).

The form of the model is illustrated in Figure 1. The population of artificial neurons that are used to represent the pronunciations of words are at the bottom of the figure, and are labelled "input units." The units that represent the spelled forms of words are at the top of the figure, labelled "output units." These two populations of units are totally interconnected *via* an intermediate layer of units called "hidden units" (so called because they are hidden from direct contact with the input or the output). The presence of

INSERT FIGURE 1 ABOUT HERE

these hidden units enables more difficult input/output mappings to be learned than would be possible if the input units were directly connected to the output units. Each input (pronunciation) unit was constrained to take on the value either 0 or 1. We wanted the model to learn to spell a vocabulary of 225 words. Thus we needed to be able to represent the pronunciation of each of those words as a pattern of 0's and 1's on the input units of the network, and have a network learn the right strengths of connection between units such that a pattern of activation would appear on the output units that corresponded to the correct spelling of that word. There were 50 input units to represent the pronunciations of words and 50 output units to represent the spelling of words. Thus a distinctive pattern of 50 0's and 1's is used to represent the spelling and pronunciation of each one of the 225 words. For each pronunciation or spelling representation, an average of 12 units were switched on and the rest were switched off. Thus "distributed representations" were employed, in that the representation of each word's pronunciation and spelling was distributed over all the input units. The input representations were chosen in such a way that words of a similar pronunciation (e.g., SOAP, HOPE) had similar representations. The spellings of words were represented in a similar way on the output units. We do not go into the full details of the representations here, (see Brown, Loosemore & Watson, 1993 for more details). However, the basic idea (following Rumelhart & McClelland, 1986b) is that each word can be seen as being composed of a series of phoneme or letter "triples." Thus, using "_" as a symbol for the space before or after a word, SOAP can be seen as composed of the four triples _SO, SOA, OAP, and AP_. If a connectionist network was given one artificial neuron for each possible triple of letters that occurs in the vocabulary to be represented, then it would be possible to represent SOAP by giving the value 1 to all the neurons that stood for one of the four triples (listed above) contained in SOAP, with every other neuron being given the value 0. This would then allow every word to be represented as a unique pattern of 0s and 1s over the set of artificial neurons.

Thus idea formed the basis of the representational scheme originally employed by Rumelhart & McClelland, 1986b, and we adopted a similar scheme in the present model althoug it differed in detail.[1]

Vocabulary of the Model

We were particularly interested in the model's ability to learn to spell three different types of word (see Table 1). The 225 word vocabulary included 19 words which were spelled in an entirely consistent way, i.e., they had only sound-to-spelling friends, where "friends" are defined as words which share the rime segment of the item (see Treiman, this volume) and are spelled in the same way). An example of such an item is the word KILL, which has only friends (HILL, WILL, TILL). In standard terminology these would be considered regular.

Word	Friends	Enemies	Classification
SOAP	none	HOPE, ROPE etc.	Irregular
KILL	HILL, WILL etc.	none	Regular
BULB	none	none	?????

Table 1

Classification of word types

words. A second set of 19 words had only sound-to-spelling "enemies," such as SOAP, which has enemies HOPE, COPE, ROPE etc. (Enemies are defined as words which share a rime pronunciation but are spelled in a different way). Such words may be considered irregular. A third category of words, like BULB, was included (BULB has no phonological neighbours spelled either the same or differently). We were particularly interested in the model's relative ability to learn to spell these different types of words with consistent or inconsistent sound-to-spelling correspondences. The remaining words in the model's 225 item vocabulary were chosen to provide friends and enemies for the words just described.

<u>Learning in the model</u>

We were interested, then, in whether the model could learn to spell words of the type we described above and, if so, which it would learn most quickly and accurately. In this session we briefly describe how the learning process works. Again only the general principles involved will be given (for details, see Brown, Loosemore & Watson, 1993).

Successful spelling in the model depends upon the network having the right strengths of connections between the units used to represent the pronunciations of words and the units used to represent the spelling of words. This is because the patterns of activation over the pronunciation units can only give rise to the appropriate patterns of activation on the spelling units if the weights on the connections between them are set appropriately. Prior to any learning, these connection strengths (weights) are set to small random values. Therefore, when a pattern of activation is imposed upon the pronunciation representing units, to represent the pronunciation of a particular word, the resulting pattern of activation that is produced on the spelling units will be random before any learning has taken place. The learning process works by imposing the pronunciation representations of each of the 225 words on the network one by one, and examining the pattern of spelling unit activation which is produced in response to each word. This, as just stated, will be random originally. For each word, the learning algorithm calculates the difference between the level of activity that is produced on each spelling unit and the level of activity that would be necessary for the correct spelling of the word to be represented. A small adjustment is made to the connectionist strength to that unit in such a way that when the same process occurs again a closer approximation to the correct pattern of spelling activation will be produced. Small incremental changes to every connection strength are made in this way so that the performance of the network gradually improves. As only incremental changes are made, it is necessary for many trials of learning to take place before any words are spelled correctly. Each "epoch" of learning consists of a presentation of all the 225 words, and an adjustment of all the connection strengths such that the spelling representations corresponding to each word's pronunciation becomes closer to the desired (correct) spelling pattern.

The performance of each word during learning can be characterised as the difference between the pattern of activity that is actually produced on the spelling units of the network when that word is presented, and the pattern of activation that would represent the correct spelling of that word. This is known as the "error score." Thus the lower the error score for a particular word at any point in learning, the better the model is doing on that word. Thus learning in the network will be reflected by a reducing error score. (There are alternative ways of assessing the performance of the network which can be interpreted more directly as a percentage of words that are spelled correctly. For a discussion of the model's performance in these terms see Brown, Loosemore & Watson, 1993).

Results

Figure 2(a) shows the average error scores during learning for 19 of each of the three word types described above. It can be seen that performance on all three word types improves over time during learning, reflected in reducing error scores. It can also be seen that words with friends and no enemies, e.g., PILL, are at all stages in learning spelled most accurately by the model, words with sound-to-spelling enemies but no friends, e.g., SOAP are spelled least accurately, and words with neither friends nor enemies come somewhere in between. An alternative way of looking at these results is to say that a given level of accuracy is achieved on consistently spelled words (those with only friends) at an earlier stage in learning than for words with sound-to-spelling enemies.

INSERT FIGURE 2 ABOUT HERE

Comparison with human data

In order to test the predictions of the model regarding the level of difficulty of the various word types, we gave four groups of children, at different stages of development, the same three sets of 19 words as had been learned by the model. To ensure that the words were in the children's vocabulary, we conducted comprehension tests on individual subjects and looked only at the error rate of the words that were known by the children. The results can be seen in Figure 2(b), which shows the error proportions for the three different word types for the four groups of children of different ages. It can be seen that the pattern of performance is very similar for the children and for the model. The children, like the model, improve in spelling accuracy over time but at all stages in development they have particular difficulty with the words with sound-to-spelling enemies (e.g., SOAP), and least difficulty with the friends-only words (e.g., PILL).

How should we interpret these results? It is important to emphasise that the good fit between model and data is achieved even though the model contains within it nothing that appears to correspond to explicit sound-to-spelling translation rules of any type. All the information in the model is included in the representations of the words and the strength of connections between the representing units. Furthermore, the mechanism of the model contains nothing which corresponds to the distinction between two different spelling routines in the standard dual route model. Thus it is possible to spell both regular and irregular items with just one mechanism.

Developmental dyslexia in the model

We wished to examine the possibility that developmentally dyslexic spelling could be characterised in terms of limited computational resources being used during learning to spell. Such an account, if it could be sustained, would contrast with traditional information processing stage accounts which refer to the absence of a particular spelling processing strategy in dyslexic populations. Seidenberg and McClelland (1989a,b; Seidenberg, 1989) have already proposed a similar account of dyslexic reading.

To investigate this hypothesis in the context of the model described above, we varied the computational capacity or resources of the network by altering the number of hidden units in the network and examining its learning performance (it will be recalled that the hidden units intervene between the input units [used to represent the pronunciation of words in the network] and the output units [used to represent the spelling of words in the network]). The larger the number of hidden units, the larger the number of connections in the network, and the greater its capacity to learn new associations. In the simulations described above, which we regard as a simulation of aspects of normal spelling development, the model was given 35 hidden units. We repeated the simulation with two further models which were identical in all respects except that one was given only 20 hidden units and the other was given only 15. These can be thought of as attempts to create a “mildly dyslexic” model and a “severely dyslexic” model.

Figure 3 shows the reducing error scores for the three types of words for all three versions for the model side-by-side. It can be seen that the “dyslexic” versions of the model learned more slowly, and reached a lower level of performance after any given amounts of learning. However all three versions of the model showed the same qualitative pattern of differences between the three different types of words throughout learning.

This is of particular interest because it has sometimes been suggested that dyslexic children will show reduced effects of spelling-to-sound or sound-to-spelling regularity (e.g. Frith, 1985). This is because one account of developmental dyslexia suggests that dyslexics have been unable to make the transition to alphabetic reading and/or spelling strategies. If this is so, they should be less liable to show effects of spelling-to-sound or sound-to-spelling regularity. Thus the presence of a regularity effect can be used as a marker for the use of alphabetic processing. If dyslexics suffer a selective problem with alphabetic processing, they should show a reduced regularity effect. In experiments designed to test this issue it is traditional to use a spelling-age match design. This involves comparing performance of dyslexic children with that of younger non-dyslexic children who are spelling at the same level. Because the two groups are matched for spelling level, it can be assumed that if any differences between the group are found this will not be simply a consequence (rather than a possible cause) of their spelling problems. We therefore carried out an analogue of the spelling-age match design on the dyslexic and non-dyslexic versions of the connectionist model. To do this, we took the three different models at the different stages in learning at which they were performing equally well on “regular” words (those with only friends). The point we chose was reached after 130 epochs of learning for the “non-dyslexic” model (with 35 hidden units), after 390 epochs for the “mildly dyslexic” model (with 20 hidden units), and after 1580 epochs for the “severely dyslexic” model (with just 15 hidden units). We then examined the error score of the models to irregular words (those with only sound-to-spelling enemies) at that point in learning. The results are shown in Figure 4(a). The error score for the three different models on regular words was the same, because the models were chosen to be the same. However, the models also showed approximately equal error scores for the irregular words. We interpret this as showing that the dyslexic and non-dyslexic models show equivalent effects of sound-to-spelling regularity.

Does the same apply to real dyslexic children? We carried out an experiment to test the predictions of the model as applied to dyslexia. We gave the same three sets of words as used in the previous experiment and simulations to two groups of dyslexic children. All the children were at a special school for children with reading problems, and on average had a reading age amore than two and a half years behind their chronological ages. In addition all children had been independently assessed as dyslexic. (For further details of the subject groups see Brown, Loosemore & Watson, 1993).

The results can be seen in Figure 4(b). This clearly shows that both groups of dyslexic subjects showed equal-sized effect of words' sound-to-spelling characteristics when compared with the non-dyslexic subjects. This is consistent with an account of dyslexia in which dyslexics eventually have the same processing strategies available to them as normal subjects, but are delayed in their acquisition of these strategies - in other words, their processing is "delayed" rather than "deviant". Work on reading in dyslexia has found similar results, in that most studies of dyslexic reading have found equivalent-sized regularity effects in dyslexic and non-dyslexics populations (Brown & Watson, 1991). However studies of reading have also shown that dyslexics tend to have particular trouble in non-word reading when compared with non-dyslexics reading at the same level (Rack, Snowling & Olson, 1992), and this has been taken as evidence for the use of qualitatively different reading strategies in the dyslexic population. Furthermore there is some evidence that a similar picture pertains in spelling, and there is evidence that dyslexics will show particular problems in spelling non-words when compared with non-dyslexics (e.g. Martlew, 1992). We therefore compared the ability of the different (dyslexic and non-dyslexic) versions of the model, again matched on spelling ability for regular words, on error score to non-words. The results are shown in Figure 5, where it is clear that the dyslexic versions of the model do show impaired non-word spelling performance, even though they have shown equal-sized effects of sound-to-spelling regularity. This would generally be taken, if found in an experiment, as evidence for qualitatively different processing in the dyslexic population. It should therefore be noted that there are no qualitatively different mechanisms in the dyslexic version of the model. Rather the models differ only quantitatively, that is, in the extent of their computational resources. Further research will be needed to confirm the non-word processing deficit in dyslexic spelling more thoroughly than has been done to date. However, it does seem that the paradoxical pattern of findings observed in dyslexic children (equal-sized regularity effects combined with a selective non-word spelling deficit when compared with appropriate controls) is reproduced in the model.

To conclude this section we suggest that much of the progression of difficulty in normal spelling, and the pattern of spelling performance exhibited by developmentally dyslexic children, can be well characterised in connectionist terms. Specifically, developmentally dyslexic spelling can be explained as reduced computational resources being available to the model.

Discussion

In this final section we review the findings, and discuss them with reference to the issues raised in the introduction to the present chapter: are computational models useful in helping us to understand spelling processes?

First of all, we focus on the ability of connectionist models to account, using very basic learning mechanisms, for data previously assumed to require the postulation of more complex cognitive mechanisms. In the most general terms, connectionist models such as the one described here operate using very simple principles of basic associative learning, of the type known to exist in animals. It is important, we suggest, to explore the ability of simple associative learning to account for so-called cognitive phenomena, for we should not invoke complex cognitive constructs such as “rules” or “strategies” in our explanations of psychological processing if we do not need to. Indeed, we see this as one of the primary roles of connectionism. Even though it may be necessary to postulate richly structured internal mechanisms to account for some aspects of language performance, it is important to account for as much data as we can in more epistemologically conservative terms. Only when we encounter data that we cannot explain in terms of simple learning principles should we be prepared to assume the existence of more complex mechanisms.

Thus we view the success of the model in accounting for much of the variance in children’s error rates, as described above, not as evidence that we have developed a complete model of spelling, but rather as evidence that some proportion of spelling error data from children can indeed be accounted for “non-cognitively.”

The approach is also of interest because it suggests a different way of viewing both normal and dyslexic spelling development - as a task of mastering the statistical associations between a set of patterns representing the phonological forms of words and a set of patterns representing the orthographic forms of words. Under this characterisation, the computational difficulty of the mapping will determine how rapidly it is learned, and this is what we found in the model and in children.

Furthermore, developmental dyslexia can be viewed as a lack of computational resources being made available in the task of learning to spell. Indeed, such an explanation can provide an account of why dyslexic children can exhibit equal-sized sound-to-spelling regularity effects as non-dyslexic children at the same time as a selective deficit in non-word processing. It is likely that degrading the quality of the phonological representations available to the model would have much the same effect, and we are currently running simulations to assess this possibility. In the case of reading, there is already plentiful evidence that the nature of the representations made available to a connectionist model can dramatically influence the performance on non-words relative to word reading. Thus the original Seidenberg & McClelland (1989a) connectionist model of reading was criticised for its poor non-reading performance (Besner, Twilley, McCann & Seergobin, 1990), but providing a similar connectionist architecture with different representations can dramatically improve reading performance (Bullinaria, 1993; Phillips & Hay, 1992; Plaut, McClelland & Seidenberg, 1992).

We suggested in the introduction that the attempt to build computational models of psychological processes can lead to the development of novel hypotheses. This was indeed the case with the model we have described briefly in this chapter: the prediction of separate “friends” and “enemies” effects, which has been confirmed experimentally, lead to the suggestion that the development of fluent and automatic application of sound-to-spelling knowledge (however that knowledge is characterised) might proceed independently of knowledge of exceptional sound-to-spelling correspondences (see Brown, Loosemore & Watson, 1993, for further details).

In conclusion, there is a long way to go before a computationally explicit model of spelling can account for all currently available spelling data. However, it could be argued that current, verbally-specified models of normal and impaired spelling development are not sufficiently explicit or even truly explanatory, for they do not specify the mechanisms underlying developmental change in enough detail to enable the models to be expressed formally. We would argue that the discipline of attempting to build working models is not only useful but essential if spelling is to be fully understood, and that the three chapters in this section of the book have all lead to insights that would have been unlikely to emerge except within a computational approach.

Footnote

[1]We used a distributed Wickeltriple representation. A complete (50 bit) pattern was created for every triple that occurred in the vocabulary. Word patterns were then made by superimposing the relevant triple patterns and assigning the ON value to any unit for which one or more of the triples had an ON value. The same method was used to construct patterns for both phoneme triples and letter triples. Each of the output units is assigned a unique set of similar triples (from the set of all possible triples, not just those that occur) for which it will tend to be in the ON state. To make each set, three groups of 15 letters are chosen, each group being associated with a particular position in the triple. The response set then comprises all those triples that can be made by selecting one letter from each of the three groups (a total of 3375). Rather than simply allow an output unit to be in the ON state if the triple occurred in that unit's response set, we ensured that all triples had exactly the same number of units in the ON state (the four whose response sets most closely matched the triple). The net effect of this encoding scheme is that, to the extent that two triples overlap, they tend to be represented by patterns with ON bits in similar positions. Although there are well-known limitations to this type of representation (Pinker & Prince, 1988), it suffices for the vocabulary of the present model.

References

BESNER, D., TWILLEY, L., McCANN, R.S. & SEERGOBIN, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review,* 97, 432-446.

BULLINARIA, J. (1993). *Neural network models of reading without wickelfeatures*. Unpublished manuscript: University of Edinburgh.

BROWN, G.D.A., LOOSEMORE, R., & WATSON, F.L. (1993). *Normal and dyslexic spelling: A connectionist approach*. Manuscript under revision.

BROWN, G.D.A. & WATSON, F.L. (1991). Reading in developmental dyslexia: A connectionist approach. In M. Snowling & M. Thomson, (Eds.), *Dyslexia: Integrating theory and practice*. Oxford: Whurr.

FRITH, U. (1985). Beneath the surface of developmental dyslexia. In K.E. Patterson, J.C. Marshall & M. Coltheart (Eds.), *Surface dyslexia*. Hillsdale, NJ: LEA.

MARTLEW, M. (1992). Handwriting and spelling: Dyslexic children's abilities compared with children of the same chronological age and young children at the same spelling level. *British Journal of Educational Psychology*, 62, 375-390.

McCLELLAND, J.L., & RUMELHART, D.E. (1986) (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2* . Cambridge, MA: MIT Press/Bradford Books.

NORRIS, D. (1993). *A quantitative model of reading aloud*. MRC Applied Psychology Unit, Cambridge: Unpublished manuscript.

PHILLIPS, W.A., & HAY, I.M. (1992). *Computational theories of reading aloud: Multi-level neural net approaches*. Technical report CCCN-13, Stirling University, December 1992.

PLAUT, D,C., McCLELLAND, J.L., & SEIDENBERG, M.S. (1992). *Reading exception words and pseudowords: Are two routes really necessary?* Paper presented at the Annual Meeting of the Psychonomic Society, St Louis, MO, November 1992.

PRINCE, A. & PINKER, S. (1988). Wickelphone ambiguity. *Cognition*, 30, 189-190.

RACK, J.P., SNOWLING, M.J., & OLSON, R.K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly,* 27(1), 29-51.

RUMELHART, D.E., & McCLELLAND, J.L. (1986a) (Eds.), *Parallel Distributed Processing: Explorations in the microstructure of cognition, Volume 1*. Cambridge, Mass.: Bradford Books/ MIT Press.

RUMELHART, D.E. & McCLELLAND, J.L. (1986b). On learning the past tenses of English verbs. In J.L. McClelland & D.E. Rumelhart (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2* . Cambridge, MA: MIT Press/Bradford Books.

SEIDENBERG, M.S. (1989) Visual word recognition and pronunciation: A computational model and its implications. In W. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, Mass: MIT Press/ Bradford Books.

SEIDENBERG, M.S. & McCLELLAND, J.L. (1989a). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523-568.

SEIDENBERG, M.S. & McCLELLAND, J.L. (1989b). Visual word recognition and pronunciation: A computational model of acquired skilled performance and dyslexia. In A. Galaburda (Ed.), *From reading to neurons*. Cambridge, Mass: MIT Press/ Bradford Books.

Figure captions

Figure One

The architecture of the connectionist model. Only some units and connections are illustrated.

Figure Two

(a) Error score to the three different word types in the model during learning

(b) Percentage errors for the three different word types
for normal subjects in four spelling ability / age groups

Figure Three

Error score to the three different word types
in “normal” and “dyslexic” versions of the model

Figure Four

(a) Error score to regular and irregular words in the three versions of the model.

(b) Percentage errors for three different word types for dyslexic and control subjects

Figure Five

Error scores to non-words and regular words in the “normal” and “dyslexic” versions of the model