London Review of Education
Vol. 4, No. 1, March 2006, pp. 5–17

Dual route and connectionist models of reading: an overview

Max Coltheart*
Macquarie University, Australia

Reading researchers seek to discover exactly what kinds of information-processing activities go on in our minds when we read; to discover what the structure and organization is of the cognitive system skilled readers have acquired from learning to read. Little is known about how the most elaborate aspects of this system work, but much has been learned about its basic building blocks such as letter identification, visual word recognition and knowledge of letter-sound rules. I contrast two approaches to theorizing about these basic reading components, the dual route approach and the connectionist approach, and offer reasons for believing that the dual route approach is to be preferred.

Introduction

Cognitive psychology views reading as an information processing activity: reading aloud is transforming print into speech, and reading comprehension is transforming print into meaning. Cognitive psychologists interested in reading seek to understand the nature of the mental information-processing systems people use to perform these transformations; and cognitive psychologists interested in learning to read seek to understand how children acquire these mental information-processing systems.

It does not seem likely that much progress would be made if we started off by investigating 'real reading', seeking for example to discover how readers, as they read *The Brothers Karamazov,* develop an understanding of what life might have been like in Imperial Russia. No one has any idea about how to carry out such an investigation; so more tractable reading situations have to be studied first. This is done by breaking up 'real reading' into simpler component parts that are more immediately amenable to investigation, with the hope that as more and more of these component

*Macquarie Centre for Cognitive Science, Macquarie University, Sydney NSW 2109, Australia. Email: max@maccs.mq.edu.au

ISSN 1474-8460 (print)/ISSN 1474-8479 (online)/06/010005–13

DOI 10.1080/13603110600574322

parts come to be understood we will get closer and closer to a full understanding of 'real reading'.

One of these component parts of reading is *visual word recognition*. When we open a novel and read

> ALEXEY Fyodorovitch Karamazov was the third son of Fyodor Pavlovitch Karamazov, a landowner well known in our district in his own day, and still remembered among us owing to his gloomy and tragic death, which happened thirteen years ago, and which I shall describe in its proper place.

we encounter many highly familiar words—*third, son, landowner, district*—which carry most of the meaning of the sentence and which we recognize immediately and effortlessly (if we are skilled readers). Part of understanding how people comprehend a whole book is understanding how they comprehend whole sentences; and part of understanding how people comprehend whole sentences is understanding how they recognize whole words. So if we knew how people recognize whole words on the page, we would know part of what we need to know in order to understand how people comprehend whole printed sentences, and even a little part of what we need to know in order to understand how people comprehend *The Brothers Karamazov.*

Consequently we need a task which measures visual word recognition and which will allow us to test theories about how this component of the reading process works; and the most frequently used task here is lexical decision. This is the task of deciding whether a printed letter string is a real word or a non-word.

Now, if I ask you which of *horse* and *zloty* is a real word of English, you don't need to consult your knowledge of whole words: the mere fact that one of these two items begins *zl* is enough to indicate that it isn't an English word, and therefore that the other letter string must be the word. So the non-word items in a lexical decision task must be word-like non-words, such as *eam, biddle* or *cloot.* There's nothing about the letter string *biddle* that rules it out as a word; the only way you can decide that *riddle* is a word and *biddle* isn't is to consult your entire list of the words you know—your *mental lexicon*—to discover that *riddle* is in that lexicon whereas *biddle* is not. Finding a printed word in your mental lexicon is exactly what we mean by visual word recognition. So if we knew how people perform the lexical decision task, that could tell us how people recognize whole words as they are doing real reading.

An obvious way in which people might perform the visual lexical decision task is simply to search through their mental lexicons one word at a time, comparing each word they find in the mental lexicon with the letter string they are looking at. If they find a match they respond 'Yes'; if they get to the end of the lexicon and have not found a matching word, they respond 'No'. But this can't be the way the task is done. Skilled language users have large vocabularies; there may be 20,000 or more words a skilled reader is familiar with in print. In lexical decision experiments, it takes people around 600 milliseconds—six tenths of a second—to respond 'No'. If responding 'No' requires checking through 20,000 words one after the other, then each check of a word is being completed in at most .03 milliseconds—3 100,000ths of a second. No one thinks the brain could possibly work that fast. So we need a different theory about how people recognize words.

I'll say more about this later. But first let us return to *The Brothers Karamazov*. It has many sentences like this one, from chapter 37:

> And when Grigory Vassilyevitch wakes up he is perfectly well after it, but Marfa Ignatyevna always has a headache from it. So, if Marfa Ignatyevna carries out her intention to-morrow, they won't hear anything and hinder Dmitri Fyodorovitch.

Who are these people? Especially, who is Marfa Ignatyevna? This is hard to figure out, because her name has only appeared once in the book so far, and that was back in chapter 14. Russian novels characteristically have many characters (*The Brothers Karamazov* has 48) and their names are not familiar to the non-Russian reader—that is, the names would not be present in the mental lexicons of such readers. Despite this unfamiliarity, the names need to be remembered if the plot is to be followed. Non-Russian readers characteristically report that they do this by pronouncing the unfamiliar names to themselves and remembering the pronunciations. So adept readers, having seen the name Marfa Ignatyevna in chapter 14, will generate a pronunciation from that unfamiliar letter sequence and store it in memory. Then when something visually unfamiliar crops up in chapter 37—the letter string Marfa Ignatyevna—the pronunciation that is generated for this will match the previously stored pronunciation and allow a connection to be made back to the events of chapter 14.

Why is it that the pronunciations rather than the visual forms of these unfamiliar letter strings are stored? Because our visual memories have far lower capacities than our phonological memories. This is easy to demonstrate. Imagine you are staying in Paris at a hotel in the Rue du Faubourg Saint-Denis. You are out walking in the Tenth Arrondissement and now you want to go back to your hotel. So you look at the street signs to try to find the street of your hotel. Do you compare each street sign to your memory of the 23 letters in the name of your hotel's street? No; no matter how execrable your French, you silently pronounce the name of each street and compare it with your memory of the pronunciation of the name of the hotel's street. The reason is that you can remember the sounds of the street name but not its letters. The same is true for Marfa Ignatyevna's name.

I have discussed this just to make a simple point: the skilled reader can translate a letter string from print to speech even if that string has never been seen before (you can read aloud letter strings such as *Ignatyevna* or *biddle*). The mental machinery that allows us to do this is therefore a second component part of the whole system we use for reading. The first part I discussed is visual word recognition, which consists of locating a familiar printed word in one's mental lexicon. This allows the reader not only to recognize a printed word, but also to then read it aloud, because part of a word's representation in the mental lexicon is a specification of how it is pronounced; and so this way of reading aloud is called the *lexical* procedure for reading aloud. When you read aloud a completely unfamiliar letter string such as *biddle*, you are not using your mental lexicon, because there is no representation of that letter string in your mental lexicon: so whatever procedure you use to translate this letter string from print to speech (of which more later) is appropriately referred to as the *non-lexical* procedure for reading aloud.

These are pretty elementary components of the act of real reading; but they are nevertheless crucial components. You won't get far with *The Brothers Karamazov* unless you are a fluent recognizer of familiar printed words; and you'll struggle with the gigantic cast of characters unless you can translate the printed name of each new character into a pronunciation (even though you've never seen this name before) to store for later reference.

Although there's a great deal about the act of reading *The Brothers Karamazov* that reading researchers don't understand, they have learned a great deal over the past thirty years about how these two elementary components of the act of real reading are actually accomplished. Figure 1 represents in diagrammatic form the idea that there's a lexical procedure for reading aloud (one that consults the mental lexicon) and also a non-lexical procedure for reading aloud (one that doesn't consult the mental lexicon). This basic idea is known as the *dual route* theory of reading aloud because it involves two routes from print to speech.

Our mental lexicons contain at least three kinds of information about words: we know about their spellings, their pronunciations and their meanings. Figure 1 includes all three types of information in a single system, the mental lexicon. This turns out, however, to be wrong; the results of research, particularly neuropsychological research with people whose language has been disturbed by brain damage, compels us to adopt the view that these three forms of information about words are stored in three separate systems, as shown in Figure 2. One of these three lexicons is the *orthographic lexicon* which represents knowledge about the visual forms—the spellings—of words. A second lexicon is the *phonological lexicon* which represents

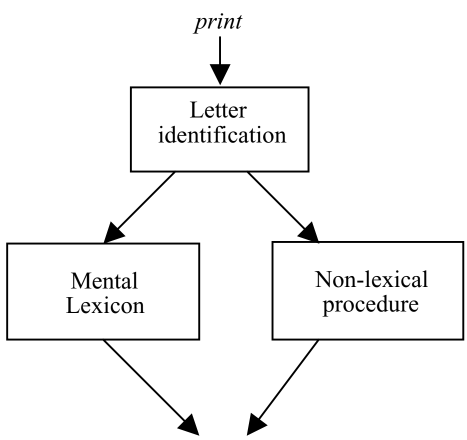

Figure 1. The dual-route theory of reading aloud

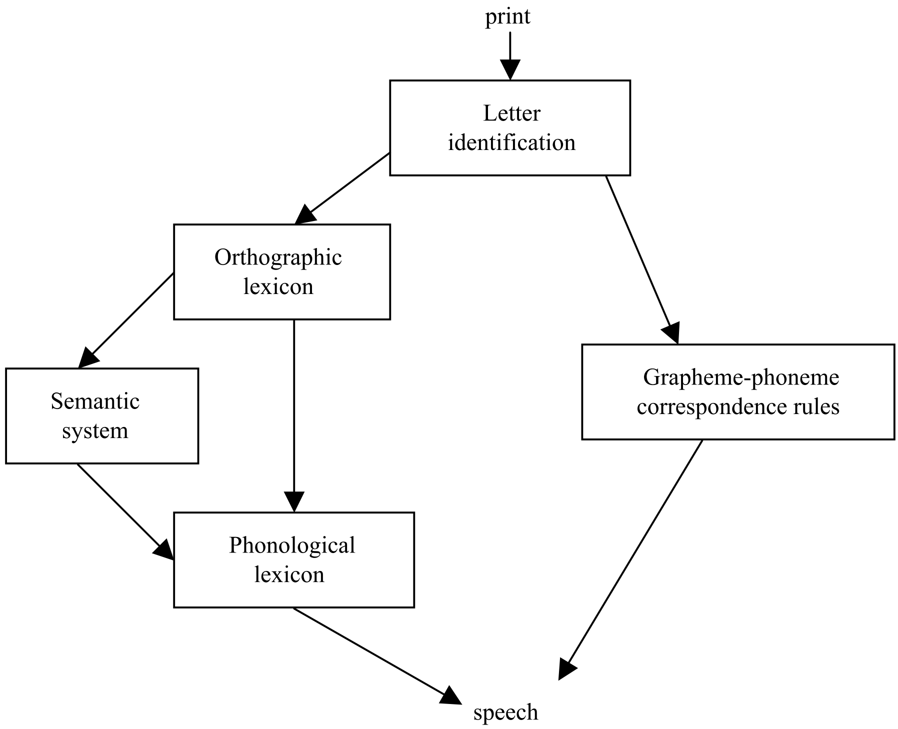

Figure 2. Elaboration of the dual-route theory of reading aloud

knowledge about the pronunciations of words. The third lexical system is the *semantic system,* where information about the meanings of words is stored.

A crucial distinction inherent in Figure 2 is between regular words and irregular words. Regular words are those that obey the grapheme–phoneme correspondence rules of English: words like *maid* or *cave.* Irregular words are those words which violate such rules: words like *said* or *have*. Regular words can be correctly read by the lexical and the non-lexical reading routes, but irregular words can be read correctly only by the lexical reading route: the non-lexical route will get them wrong (it will read *said* to rhyme with 'maid', *have* to rhyme with 'cave'—and *yacht* to rhyme with 'matched').

It is possible to show that in people with different forms of brain damage, any two of these lexical systems can be intact while the third is damaged.

For example, in some people with dementia (Blazely *et al*., 2005), knowledge of word meanings is severely impaired, but the person with dementia can still perform the visual lexical decision task with normal accuracy (so the orthographic lexicon is still intact) and can still say words with normal accuracy (so the phonological lexicon is still intact); here the only one of the three lexical systems that is impaired is the semantic system. In people with the form of aphasia known as *anomia* (for review see Nickels, 1997), it is very difficult to access the pronunciations of words in the phonological lexicon, but there can be normal visual word recognition and normal knowledge of word meanings; here the only one of the three lexical systems that is impaired is the phonological lexicon. And, finally, in the form of acquired reading disorder known as *surface dyslexia* (Patterson *et al*., 1985), the affected person can still see perfectly well but can no longer recognize many formerly familiar printed words, even though still able speak those words and still able to appreciate their meanings when the words are heard; here the only lexical system that is impaired is the orthographic lexicon. Such a person may well read *said* to rhyme with 'maid', *have* to rhyme with 'cave', and *yacht* to rhyme with 'matched', because the lexical route can no longer be used to read these words. Regular words and non-words will still be read perfectly because the intact non-lexical route can get them right.

It is hard to see how these various forms of neuropsychological impairment could be explained unless the mental lexicon is broken up into these three separate systems as proposed in Figure 2.

A second way in which the elaborated dual route model in Figure 2 differs from the basic model in Figure 1 is that it makes a specific proposal about how the non-lexical procedure for reading aloud works, namely, that when people read aloud without making reference to the mental lexicon, they do this by applying their knowledge of grapheme-to-phoneme correspondence rules.[1]

If Figure 2 is a correct account of basic parts of the reading system of skilled readers, then when children are learning to read they are going to have to progress towards acquiring the mental architecture depicted in Figure 2 if they are to become skilled readers. That is, possession of this mental processing architecture is an end-product of successful reading acquisition. There are two different ways of thinking about reading acquisition in this context. *Stage* theories of reading acquisition (see Frith, 1985) propose that as children learn to read they pass through a series of stages that involve qualitatively different ways of reading. On this view, the Figure 2 diagram does not describe how children read until they have reached the final stage, i.e., fully skilled reading. An alternative view (see Marshall, 1984; Jackson & Coltheart, 2001) is that children at various different points in the course of learning to read an alphabetic language differ only quantitatively: after they have begun to learn to read, they all have orthographic lexicons, for example, and differ only in how many words are represented in these lexicons; they all have a non-lexical procedure for reading aloud, for example, and differ only in the scope and sophistication of the knowledge of the relationships between letters and sounds that can be used by that procedure. On this view, the Figure 2 diagram does describe how children read even when they are only beginning readers; learning to read does not involve a progression through qualitatively different stages but instead a progressive quantitative expansion of the system shown in Figure 2. I consider that the evidence favours this latter view over a stages view; for further discussion of this see Marshall (1984), Jackson & Coltheart (2001), Coltheart (2006) and Castles *et al*. (2006).

One can also apply this particular way of thinking about normal reading acquisition from the perspective of Figure 2 to the question of abnormal reading acquisition—that is, developmental dyslexia.[2] According to Figure 2, if a child is to be normal-for-age at reading, then each of the four components of the system in Figure 2 will have to have been acquired to an age-appropriate level. If any one of the components has not been acquired to that level, then reading will not be normal-for-age; and the way in which reading will be abnormal for age will depend on which component or components of the system has not been acquired normally. Suppose just the semantic system is developmentally impaired: such a child will be able to read aloud normally but will have impaired reading comprehension (and impaired comprehension of spoken language). This does occur in some children, and is known as *hyperlexia* (see Aram, 1997). If just the orthographic lexicon is poor for age, the child will be normal at reading regular words and non-words, but poor at reading irregular words; this is *developmental surface dyslexia*—see Castles and Coltheart (1993, 1996) and for review Jackson and Coltheart (2001) and McDougall *et al.* (2005). If just the non-lexical procedure for reading is poor for age, the child will be normal at reading regular words and irregular words, but poor at reading non-words; this is *developmental phonological dyslexia*—see Castles and Coltheart (1993) and Howard and Best (1996), and for review Jackson and Coltheart (2001) and McDougall *et al*. (2005).

Understanding how children learn to read involves seeking answers to two questions. The first is: *what is it that the children are learning*? The second is: *how are they learning it*? Figure 2 proposes an answer to the first question: what children are learning as they learn to read is the dual route mental architecture depicted in Figure 2. But dual route theorists have said little about how children learn this architecture: little about, for example, how new representations are introduced into the orthographic lexicon, or how the child learns letter-sound rules. A different approach to theorizing about the reading system, the connectionist approach, does say something about the learning of reading; perhaps for this reason the connectionist approach is the superior approach?

Connectionist modeling of reading

This approach (see Plaut *et al*., 1996; Harm & Seidenberg, 1999, 2004) differs from the dual route approach in three fundamental ways, as follows.

- *The nature of representation: local versus distributed.* According to the dual route approach, words are represented *locally* in the reading system. What this means is that a word corresponds to a single unit in the lexicon, a single lexical entry. The contrasting idea is *distributed representation*, as used in connectionist models: any word is represented by the activation of numerous units in the system, and any unit in the system plays a role in representing many different words.
- *The nature of processing: parallel versus serial*. In all the connectionist models of reading, all processing goes on in parallel: for example, when a non-word is presented to the model, all the letters are processed by the model simultaneously. In contrast, one component of the dual route model in Figure 2 operates serially: the non-lexical procedure translates letters to sounds one letter after another, from left to right.

- *Learning*: the knowledge connectionist models use to carry out the reading-aloud task is not provided by the modeller, but developed over time by the model as it is repeatedly exposed to the spellings of words and their correct pronunciations—the knowledge is developed under the control of a learning algorithm which gradually and progressively adjusts the strengths of the connections in the model so that the model's response to each word in its training set becomes progressively more accurate. In contrast, while the knowledge that a dual route model uses to carry out the reading-aloud task is assumed to be gradually acquired by children, the model is intended just as a description of the information-processing system that children acquire as a result of this learning.

We can thus directly compare the two approaches by considering three questions, as follows.

Are the representations in the reading system local or distributed?

The key issue here is lexical decision. Skilled readers find it very easy to decide rapidly that *riddle* is a real word and *biddle* is not. How do they do this? Models that propose that representations in the reading system are local offer a simple answer: any kind of interrogation of the orthographic lexicon will show that there's a unit in that system corresponding to the word *riddle* but no unit corresponding to the non-word *biddle*, and so responding 'Yes' to *riddle* and 'No' to *biddle* is simply done. Serial checking of one item after another in the mental lexicon – which can't be the way lexical entries are found, as noted earlier—is not needed: in dual route models of reading the right lexical entry can be directly activated. If the letter *c* is connected to all words beginning with *c*, the letter *a* connected to all words with an *a* in the second position, and the letter *t* connected to all words with a *t* in the third position, then the letter string *cat* will activate the right entry in the orthographic lexicon without any need for a search through the lexicon. That is how the dual route model of reading works (Coltheart *et al.*, 2001).

But according to connectionist models, since there are no local representations, there's no unit in the system uniquely responsive to *riddle*: the word *riddle* will activate a lot of units (since its representation is distributed) but so will the non-word *biddle*. Therefore:

> … given that the current distributed approach to lexical processing does not contain word-specific representations, it becomes important to establish that distributed models can, in fact, perform lexical decision accurately and that, in doing so, they are influenced by properties of the words and non-words in the same way as human readers. (Plaut, 1997, p. 785)

Advocates of connectionist modeling of reading have not succeeded in establishing this. There has been no report of a connectionist model of reading that can perform the lexical decision task accurately and which in so doing is influenced by the various properties of words and non-words that are known to affect human performance of the lexical decision task. The only work here has sought to show that, if a connectionist model is equipped with a simulated semantic system, it can discriminate words from non-words because words activate the simulated semantic system more strongly (Plaut, 1997; Harm & Seidenberg, 2004). If the correct explanation of how human readers perform the lexical decision task is that they rely on consulting their semantic systems, then people whose brain damage has left them with severe semantic impairments could never achieve normal accuracy in the visual lexical decision task; but some can (for examples, see Lambon Ralph *et al.*, 1995, 1996, 1998; Ward *et al.*, 2000; Blazely *et al.*, 2005).

Is all the processing performed by the reading system parallel processing, or are some reading operations performed in a serial manner?

Rastle and Coltheart (2006) have reviewed this literature. They identify eight different effects reported in the literature on reading aloud which cannot be explained unless there is some form of serial processing occurring in the reading system. I'll give just two examples.

- *Position of irregularity:* skilled readers are slower at reading aloud irregular words than regular words. The size of this regularity effect depends upon where in the irregular word the irregular grapheme–phoneme correspondence is. Rastle and Coltheart (1999) showed that the effect was large when the irregularity involves the first grapheme of an irregular word (as in the word *chef*), smaller but still significant when it involves the second grapheme (as in the word *sown*), and absent when it involves the third (as in the word *steak*). This was confirmed by Roberts *et al.* (2003), who also showed that this positional effect could not be simulated by a connectionist model of reading which had no serial-processing component but could be correctly simulated by the Dual Route Cascaded (DRC) model (Coltheart *et al.*, 2001) of reading, a computational realization of the dual route theory shown in Figure 2, because, in that model, the non-lexical procedure operates from left to right.
- *Position-sensitive Stroop effect*: when your task is to name the colour in which a word is printed, your colour-naming response is faster when the colour name you have to produce starts with the same phoneme as the word that is carrying the colour: if the colour name you have to utter is 'red' you are faster when the word is *rat* than when it is *cot*. There's also facilitation from shared final phonemes: the response 'red' is faster to the word *bad* than to the word *cot.* But facilitation is larger when it is the first phoneme that is shared than when it is the last (Coltheart *et al.,* 1999). This also indicates that there is some serial processing going on as we read, and this effect too can be simulated by the DRC model of reading in which the non-lexical procedure operates from left to right. It does not seem that connectionist models (in which all processing is parallel) could ever explain this position-sensitive effect.

Do connectionist models offer a satisfactory account of how children learn to read?

When a connectionist model can *learn* to read via repeated exposures to printed words and their correct pronunciations, that's clearly a fascinating feature of the model. But such a model could only be preferred to dual route models, which do not offer any account of how reading is learned, if the connectionist model learns in the same way that children learn. If connectionist models do not learn in the same way as children do, then the fact that these models learn offers no reason to favour them over models that do not learn. So a key question is: could the way connectionist models learn to read be the same as the way that children learn to read? I will argue that the answer to this question is no, for two reasons.

Firstly, the number of exposures to spelling-sound pairs that are needed for the connectionist models to learn them seems vastly greater than what is needed by children. There are numerous different connectionist models of learning to read, all somewhat different—Coltheart (2005) lists seven such models—but the following points apply to all of them. The first of the connectionist models of reading described in Plaut *et al.* (1996) was trained on a set of 2998 words. Each word and its pronunciation was presented 300 times to the model; after that number of presentations, the model could correctly generate the pronunciation of all the words in the training set. Another of the connectionist models described in Plaut *et al.* (1996) needed to have each word presented 1900 times and even after that had not learned the pronunciations of 25 of the words in the training set. Harm and Seidenberg (1999) used a training set of 3123 words for their connectionist model, and training consisted of presenting a total of ten million words. Harm and Seidenberg (2004) used a training set of 6103 words and their connectionist model was trained for 1.5 million word presentations. It is clear from these figures that the PDP models require a word to be presented along with its correct pronunciation hundreds or thousands of times for its pronunciation to be learned correctly. Children learning to read aloud single words do not require to see each word and hear its pronunciation hundreds or thousands of times before they can correctly read aloud the words.

A possible defence to this criticism is that children are not expected to learn to read some thousands of words at the same time: perhaps if they were, they might require hundreds or thousands of exposures to each word (though this seems very unlikely). Instead, the more usual scenario is that children learn a small set of words until they have got them right, then another small set and so on. But this introduces the second and even more serious problem for learning in connectionist models. Suppose you train a connectionist model on some small set of words—call it Set A—until all are read perfectly. Then you train the model to perfection on a new set, Set B, without further training on Set A. If you now retest it on Set A, it will perform very poorly. This is the problem of *catastrophic forgetting* in connectionist networks (McCloskey & Cohen, 1989). It occurs because the connection strengths that are learned so as to be able to get all the Set A words correct are all liable to be changed when the only current task is to get all the Set B words right: the Set A words no longer matter to the model now. The only way to avoid the catastrophic forgetting problem here is to keep presenting the A words and their correct pronunciations even after training on the B words has begun. But nothing like this happens as humans learn to read. Once we have learned to read a word aloud, we do not need to keep receiving training on what its correct pronunciation is; we do not unlearn it as soon as that training is discontinued. Of course, we keep *seeing* these words over and over again as our lives progress, and we also keep hearing them over and over again. But we do not need constant feedback regarding a printed word's pronunciation; once the word is learned, it stays learned even if we never again are taught it by being shown it as we are told how to say it. In contrast, when PDP networks are given a new set of words to learn and training on a previously learned set is discontinued, the old set will be forgotten. That doesn't happen with children.

My conclusion then regarding the connectionist approach to explaining reading is that at present it has not been successful. The connectionist eschewal of local representations makes it impossible for current connectionist models of reading to explain how people perform the lexical decision task. The connectionist insistence on parallel processing makes it impossible for current connectionist models of reading to explain the many demonstrations of serial processing as people read aloud. And though connectionist models do learn to read, this is not an advantage of such models, because they don't learn to read the way children do. So at present the dual route approach described in Figure 2, and converted by Coltheart *et al.* (2001) into a computational model of visual word recognition and reading aloud (the DRC model), seems to offer the most promising initial steps down the long path towards an understanding of what readers are doing as they are understanding *The Brothers Karamazov*.

Acknowledgements

The author thanks Genevieve McArthur, Morag Stuart and Taeko Wydell for helpful comments on an initial draft.

Notes

1. The technical term 'grapheme' refers to the written representation of a phoneme. So, for example, the word *sheep* has five letters but only three graphemes, these graphemes being *sh*, *ee* and *p*.
2. By 'developmental dyslexia' I mean simply difficulty in learning to read. This contrasts with 'acquired dyslexia' which refers to any impairment of reading that is caused by brain damage in a person who had learned to read normally prior to suffering that brain damage.

Notes on contributor

Max Coltheart is Scientific Director of the Macquarie Centre for Cognitive Science and James Packer Professor of Educational Research. He's held posts at universities in Australia, Canada, Italy, the US and England, is author of 229 journal

articles and book chapters, and has co-authored the books *Language processing in children and adults* and *Routes to reading success and failure*.

References

Aram, D. M. (1997) Hyperlexia: reading without meaning in young children, *Topics in Language Disorders,* 17, 1–13.

Blazely, A., Coltheart, M. & Casey, B. (2005) Semantic dementia with and without surface dyslexia, *Cognitive Neuropsychology,* 22, 695–717.

Castles, A., Bates, T. & Coltheart, M. (2006) John Marshall and the developmental dyslexias, *Aphasiology.*

Castles, A. & Coltheart, M. (1993) Varieties of developmental dyslexia, *Cognition,* 47, 149–180.

Castles, A. & Coltheart, M. (1996) Cognitive correlates of developmental surface dyslexia: a single case study, *Cognitive Neuropsychology,* 13, 25–50.

Coltheart, M. (2005) Modelling reading: the dual route approach, in: M. J. Snowling & C. Hulme (Eds) *The science of reading* (Oxford, Blackwells Publishing).

Coltheart, M. (2006) John Marshall and the cognitive neuropsychology of reading, *Cortex.*

Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001) DRC: a dual route cascaded model of visual word recognition and reading aloud, *Psychological Review,* 108, 204–256.

Coltheart, M., Woollams, A., Kinoshita, S. & Perry, C. (1999) A position-sensitive Stroop effect: further evidence for a left-to-right component in print-to-speech conversion, *Psychonomic Bulletin and Review,* 6, 456–463.

Frith, U. (1985) Beneath the surface of developmental dyslexia, in: K. E. Patterson, J. C. Marshall & M. Coltheart (Eds) *Surface dyslexia: cognitive and neuropsychological studies of phonological reading* (Hove, Lawrence Erlbaum Associates).

Harm, M. & Seidenberg, M. S. (1999) Reading acquisition, phonology, and dyslexia: insights from a connectionist model, *Psychological Review,* 106, 491–528.

Harm, M. & Seidenberg, M. S. (2004) Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes, *Psychological Review,* 111, 662–720.

Howard, D. & Best, W. (1996) Developmental phonological dyslexia: real word reading can be completely normal, *Cognitive Neuropsychology,* 13, 887–934.

Jackson, N. & Coltheart, M. (2001) *Routes to reading success and failure* (Hove, Psychology Press).

Lambon Ralph, M. A., Ellis, A. W. & Sage K. (1998) Word meaning blindness revisited, *Cognitive Neuropsychology,* 15, 389–400.

Lambon Ralph, M. A., Sage, K. & Ellis, A. W. (1996) Word meaning blindness: a new form of acquired dyslexia, *Cognitive Neuropsychology,* 13, 617–639.

McCloskey M. & Cohen, N. J. (1989) Catastrophic interference in connectionist networks: the sequential learning problem, in: G. H. Bower (Ed.) *The psychology of learning and motivation* (New York, Academic Press).

McDougall, P., Borowsky, R., MacKinnon, G. E. & Hymel, S. (2005) Process dissociation of sight vocabulary and phonetic decoding in reading: a new perspective on surface and phonological dyslexias, *Brain & Language,* 92, 185–203.

Marshall, J. C. (1984) Toward a rational taxonomy of the developmental dyslexias, in: R. N. Malatesha & H. A. Whitaker (Eds) *Dyslexia: a global issue* (The Hague, Martinus Nijhoff).

Nickels, L. A. (1997) *Words fail me: spoken word production and its breakdown in aphasia* (Hove, Psychology Press).

Patterson, K. E., Marshall, J. C. & Coltheart, M. (Eds) (1985) *Surface dyslexia: cognitive and neuropsychological studies of phonological reading* (Hove, Lawrence Erlbaum Associates).

Plaut, D. C. (1997) Structure and function in the lexical system: insights from distributed models of word reading and lexical decision, *Language & Cognitive Processes,* 12, 767–808.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S. & Patterson, K. (1996) Understanding normal and impaired word reading: computational principles in quasi-regular domains, *Psychological Review,* 103, 56–115.
Rastle, K. & Coltheart, M. (1999) Serial and strategic effects in reading aloud, *Journal of Experimental Psychology: Human Perception & Performance,* 25, 482–503.
Rastle, K. & Coltheart, M. (2006) Is there serial processing in the reading system; and are there local representations?, in: S. Andrews (Ed.) *All about words: current issues in lexical processing* (Hove, Psychology Press).
Roberts, M., Rastle, K., Coltheart, M. & Besner, D. (2003) When parallel processing in visual word recognition is not enough: new evidence from naming, *Psychonomic Bulletin & Review,* 12, 405–414.
Ward, J., Stott, R. & Parkin, A. J. (2000) The role of semantics in reading and spelling: evidence for the 'summation hypothesis', *Neuropsychologia,* 38, 1643–1653.