When is a Developmental Model not a Developmental Model?

Steve Cassidy

Published in: Cognitive Systems, 2-4, 1990

Abstract

A recent paper by Seidenberg and McClelland (1989) describes a computational model of word recognition and naming. The authors claim that it is a developmental model; that is, it explains how word recognition skills are acquired by children. The purpose of this paper is to challenge that claim and, in doing so, to set out a number of criteria against which a model of reading development can be assessed.

To summarize the criteria:

- The environment of learning should reflect that of the child.
- The representations used should be accurate and adequate for learning.
- The model's performance should reproduce observations of children's performance.
- The model should be consistent with wider theories of cognition.

Some criticisms of the Seidenberg and McClelland model arise directly from their choice of a connectionist architecture for implementing the model. We try to identify those parts of the model that result from this choice and argue that this type of connectionist model is an inappropriate way of describing learning.

We claim that this discussion of models of word recognition is also relevant to other areas of cognitive development.

COGNITIVE DEVELOPMENTAL MODELS

The construction of computational models is an important adjunct to theory building in cognitive science. Such an exercise leads to a better understanding of the detail of the processes involved in cognition by forcing the theorist to consider the fine grain, computational implications of the theory. Another reason for building an explicit model is to find out the behavior predicted by complex theories in realistic situations;

it is often not possible to predict what will happen in response to complex stimuli without running a computer simulation. This is particularly true of developmental theories, where the learning environment provides a large and complex stimulus set.

The psychology of word recognition is rich in alternative theories; computational modelling can help to evaluate the alternatives. For example, many theorists have put forward the idea of using analogy to recognise words (Glushko, 1979). Analogy is a very loose term that covers almost any means of accessing a lexical item from its orthography. If such theories are to be developed, the means by which analogies are made and used need to be further specified. In building a computational model we are forced to define what we mean by analogy; the performance of the model then serves as a basis for assessment of our definition.

When we look at theories of the development of reading skills, there is not as much variety as for mature reading. However, an important part of any theory of mature performance is how the mechanisms and structures it proposes develop; some do not have any obvious developmental route, which seriously harms their credibility (for example, the dual route theory (Humphreys & Evett, 1985)). A number of developmental theories (Frith, 1985; Marsh, Morton, Welch, & Desberg, 1980; Seymour, 1984) suggest that there are stages in the development of reading skills where different methods are used to decode words. These theories seems to account for a number of observations, but it is not clear that it is feasible for children to learn in this way, given the orthographic environment that they are exposed to. Building a computational model would serve to further specify the processes involved and, if successful, show that learning is possible in this way.

Seidenberg and McClelland (1989) present a "distributed, developmental model of word recognition and naming". Their model learns to associate the spelling of 2897 monosyllabic words with their pronunciations, using a connectionist learning algorithm. Once it has learned these associations, the model reproduces a number of observations of reading performance. They also observe the behavior of the model as it learns the associations, and draw conclusions about the nature of reading development.

This paper is an attempt at establishing some ground rules for the construction of a developmental model of word recognition. In order to do this we critique the model of Seidenberg and McClelland from the point of view of development, and set out some criteria that should be met by a developmental model. Although we will be dealing exclusively with word recognition in this paper, much of the discussion is relevant to other aspects of cognition.

EVALUATING A DEVELOPMENTAL MODEL

A developmental model is a model of a cognitive task that claims to give an account of how the task is learned. The particular model under consideration here is not primarily a developmental model, rather it claims to model mature reading performance. However, Seidenberg and McClelland do make some claims about development in the model and it is this aspect of their work that we examine here.

We will not discuss the particular findings that a developmental theory of reading should account for, except where they illustrate a particular point; instead we will consider methodological issues in the construction of a theory/model of the development of word recognition. The criteria we develop can be summarised under four headings:

- The environment of learning.
- The adequacy and accuracy of the representations used.
- The performance of the model, compared with empirical results.
- The relationship of the model to wider theories of cognition.

The following sections discuss the model of Seidenberg and McClelland in terms of these categories.

An Outline of the Model

SaMc[1] is implemented in a connectionist network with a set of 400 orthographic units, 200 hidden units and 460 phonological units (Figure 1). Both the orthographic and phonological units use an encoding system similar to that used in Rumelhart and McClelland's (1987) model of the acquisition of verb endings: a word is represented by the set of letter triples it contains. The representation is distributed in that each orthographic unit represents a number of letter triples, and each triple is represented in a number of units. When a word is presented to the model, the units corresponding to the letter triples in that word are activated. For example, the word CART would be represented by turning on the units corresponding to the triples *[ART, CAR, RT#, #CA]* (where # is a word boundary), and turning all other orthographic units off. Phonology is represented in a similar way, each unit corresponds to a triple of phonetic features – for example, the triple *[vowel, fricative, stop]* which is present in words like SOFT and POST.

[1] It is often necessary to refer to "Seidenberg and McClelland's model" in this discussion, rather than use that clumsy and long-winded noun phrase we have given the model a name: . We hope that this does not offend 's authors, or our readers.

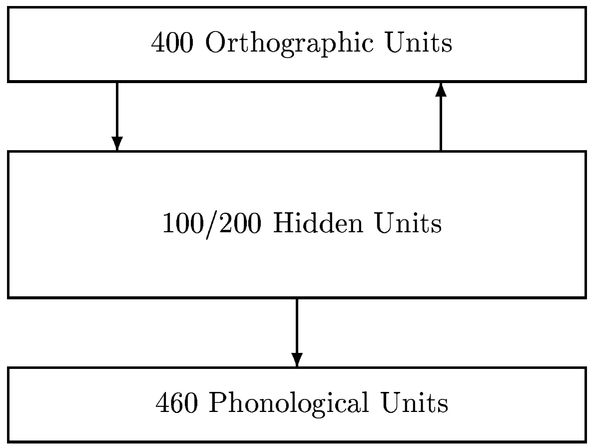

Figure 1. The organisation of SaMc

The orthographic and phonological units are connected to 200 'hidden' units. The weights on the links between the units are altered according to the back-propagation learning algorithm (Hinton, 1987) so that the network learns to associate the orthographic and phonological representations of the words in the training vocabulary.

Environment

Nothing is learned in isolation – there is always an environment that influences and aids learning. It follows then, that an important part of modelling the development of any cognitive skill is to map out the experiences which form the environment for learning. In reading, these experiences are the stories and books that the child reads, and any coaching given for particular skills, such as decoding novel words.

In a study of the vocabulary of first year reading books in New Zealand, Thompson (1982) identified 142 *types*[2] occurring more than 10 times in a 10,903 *token* sample[3]. In the books used for the first two years the vocabulary is extended to 848 types[4]. In the latter case, type frequencies range from 1 to 5470 (for *the*) with many words having a frequency of around 1000. Approximately 30% of the Thompson vocabulary are multisyllabic words.

SaMc is trained with 2897 monosyllabic words, taken from the Kucera and Francis

[2]Here we distinguish between *types*, distinct words, and *tokens*, instances of words in a text. Elsewhere we will use *word* to refer to both of these.

[3]These 142 types account for 80% of the tokens, the total number of types was 563.

[4]These 848 types account for 85% of the 88386 tokens, there were a total 3964 types.

(1967) corpus of American English. The model is trained to associate orthography-phonology pairs for 250 epochs. An epoch consists of a set of presentations of the words in the vocabulary; each word has a chance of being presented that is proportional to its frequency in the corpus. In the course of the training regime, each word is presented between about 12 and 230 times, depending on its frequency. Seidenberg and McClelland comment:

> "The sampling method is not intended to be faithful to the experience of children learning to read in American culture. In the model, all words are available for sampling throughout training, with frequency modeled by the probability of being selected on a given learning trial. In actual experience, however, frequency derives in part from age of exposure; words that are higher frequency for adults tend to be introduced earlier than lower frequency words. In learning to read then, words are introduced sequentially and often in groups to emphasize salient aspects of the orthography"

and later:

> "The model works as well as it does because it is trained on a significant fragment of written English, which contains complex latent structure."

Two points can be made about the training environment of SaMc. Firstly, consider the way in which the vocabulary is presented to the learner. In SaMc, all of the words are presented at once so that the 'statistical properties' of the vocabulary can be learned. On the other hand, a child's vocabulary grows slowly as new stories are read and new words encountered. There is no point in the training of SaMc where it can read, say, 50 words well, and 50 more with some other (contextual) clues: it always reads 2897 words with varying accuracy. Seidenberg and McClelland's assumption that learning to read is equivalent to learning about the statistical properties of the written language may be *incompatible* with a gradually expanding the reading vocabulary. It is not until the child can read a substantial number of words (say, 500) that statistical properties will be evident; this leaves at least a year of reading development unaccounted for. Van Lehn (1985) argues for felicity conditions for human learning – that the way in which material is presented can enhance its learnability. If this is the case for reading, it may be *necessary* for the vocabulary to grow slowly; Seidenberg and McClelland's model cannot address this issue. Current connectionist learning algorithms (Hinton, 1987) require that the patterns to be learned are presented as one set of stimuli; any regularities in this set will be utilized in learning associations. The use of such an algorithm forces Seidenberg and McClelland to present a large vocabulary as an homogeneous set of words rather than in a more realistic manner.

Secondly, the training vocabulary used for SaMc differs from that of a child in two important ways. The vocabulary is significantly smaller and contains no multisyllabic words. The ability of SaMc to learn depends on the presence of statistically significant regularities in the training samples; generalisations that can be made from the Kucera and Francis corpus may not be possible in a smaller vocabulary, such as that collected by Thompson (1982). By excluding multi-syllabic words from the vocabulary, the task of learning orthography/phonology association is greatly simplified; all letter clusters that occur frequently will correspond to a phoneme or phoneme cluster. Multi-syllabic words introduce ambiguity: letter clusters that occur frequently on syllabic boundaries (for example *nd* in *boundaries*) do not map onto phonemes.

Both of the above criticisms of SaMc relate to assumptions made about the training environment that may be too strong; assumptions which, we think, are in part forced by the connectionist implementation. Another observation concerns the nature of the stimulus that is available to the model, and how such stimulus is used by the model for learning.

Seidenberg and McClelland assume that there is a source of phonological input that enables SaMc to learn the association between orthography and phonology:

> "...we assume that the phonological pattern may be supplied as explicit external teaching input – as in the case where the child sees a letter string and hears the teacher say its correct pronunciation – or self-generated on the basis of the child's prior knowledge of the pronunciation of words."

If input from the teacher can only provide part of the required input then the rest must come from a source *outside the model.* We can identify a number of types of stimulus that might be available to the child learning to read (Table 1). SaMc will only learn

Table 1. Types of reading experience.

Sees	**Reads**	**Hears**	**Comment**
bat		/bæt/	Child is read to by teacher
bat		/bæd/	Child mis-hears
bat	/bæt/	'correct'	Feedback on child's reading
bat	/bæd/	'wrong'	
bt	/bæt/		Child reads her own incorrect spelling
bat			Word seen with no other stimulus

in the first two cases listed here, where orthographic and phonological stimuli are presented simultaneously. The case where the child's own reading is being corrected seems to involve a more complex process than the positive reenforcement of correct answers implied by SaMc. An interesting case is the interaction between spelling and reading – if children read their own spellings then they are likely to experience many invalid letter strings. Depending on the view taken of the relationship between reading and spelling, this may serve to either reinforce or correct an incorrectly learned orthographic representation.

The observations that Seidenberg and McClelland make of their model as it learns, and as a mature model, suggest that the mechanisms they use may be similar to those used in word recognition. However, the model is unconvincing because of the lack of attention paid to the environment in which it learns, and the degree to which its environment is shaped by the requirements of the back-propagation learning algorithm.

Representation

A cognitive model makes use of a representation of the world. The representation characterises those aspects of the world that are important to the task in a way that facilitates the task. In a model of reading, the aspects that must be represented are the orthographic and phonological forms of the words in the vocabulary, along with any semantic or other information used in recognising words. Obviously, there are many choices for this representation; the particular choice has an enormous effect on the shape of the resulting model.

Seidenberg and McClelland's choice of a connectionist architecture for the implementation of SaMc restricts the types of representations that are available to them. The model uses a representation which has been criticised elsewhere (Pinker & Prince, 1988) as being inadequate for language learning. The triple representation is unable to distinguish some possible letter strings (such as the words *algal* and *algalgal* of the Australian language Oykangand) and does not provide a natural base from which to learn about word recognition. For instance, suffixes and affixes seem to be useful 'units' that should be recognised in an input word. The triple representation makes it difficult to describe these morphemes (meaning carrying units) in a way that would facilitate the appropriate generalisations. Seidenberg and McClelland would counter this argument by claiming that the morphemes exist only as a convenient abstraction of the real representation: they are emergent from the behavior of the underlying network. It is difficult to judge the validity of such a claim, as the current model only deals with monosyllabic words which do not show a rich morphemic structure.

The problems facing SaMc are due to the difficulty of representing a *sequence* of entities in a connectionist model. Such a model can naturally represent the presence or absence of a particular feature in the input, but has difficulty representing relationships between input features – such as the spatial relationships between letters in a word. There are various ways of getting around this problem: the triple representation is one example, another is to have 26 units for each letter position in a seven-letter window (this method is used in the NETtalk text-to-speech system

(Sejnowski & Rosenberg, 1986)). Recent work is investigating the use of time delay networks for representing sequences (Elman, 1988); this is a new approach to representation in neural networks and may provide an answer to these problems. The representation of sequence in SaMc and similar networks is not adequate for modeling the development of cognitive processes in language.

Studies of the reading performance of young children have shown that, at least in the early stages, the initial and final letters are used to identify a word (Campbell, 1987). It is clear why this might be the case: the boundary letters of a word are more easily picked out visually than internal letters. In SaMc, initial and final letters have no special significance; they are just as easily *perceived* as the internal letters (this is also true of the NETtalk representation) due to the lack of an explicit representation of ordering. The model is unlikely to learn to use these salient letters as clues to word identification, unless they happen to be statistically useful in making the discrimination.

Consider also the observations made by Seymour and Elder in their longitudinal study of beginning readers (Seymour & Elder, 1986):

> "It appeared that at a certain point in development the *position* of salient letters was not used as a discriminatory feature. For example the letter "k" was a salient feature for identification of the word "black". In the 16th week of schooling, LBH responded "black" to *likes, think* and *thank*, and also to the non-word targets *bkacl, eadhk, pjoek* and *htoek.*" (Seymour & Elder, 1986)

It would certainly be possible for a connectionist model to to learn $k \rightarrow$ /blæk/; it would need to ignore the position of letters in the input word and just fix on a useful feature (letter or letter group) to identify the word. However, a child will later make the representation more precise, by making use of positional information. Our connectionist model, on the other hand, would not then be able to later take note of the position of letters, this would require it to undo a generalisation it had already made and make a different one; connectionist models are not (yet) capable of such behavior. For SaMc to learn that $k \rightarrow$ /blæk/ it would have to see a number of examples where the orthographic units corresponding to k (in a number of contexts) are active at the same time as the phonological units representing /blæk/. There is no realistic way that SaMc will learn this association.

Many authors have noted *stages* in the development of reading skills. Frith (1985) outlines three stages: an early 'logographic' stage where words are recognised on a visual basis, an 'alphabetic' stage where recognition is mediated by phonology and an 'orthographic' stage where morphemic analysis is developed. If these stages are real then they imply that different strategies are being used at different times for word recognition, which in turn implies changing requirements on the representation language. The representation language must cope with both the initial and mature versions of the representation of the lexicon and all the stages in between. Here the connectionist network is interesting in that at an early stage it will learn associations by rote, and then, when generalisations have been made, progress to a stage that resembles the use of letter to sound rules. Thus a connectionist model is able to simulate two distinct 'strategies' for recognition using a single underlying representation and recognition procedure.

The issue of representation becomes important when a computational account of a cognitive task is developed. Clues as to the nature of the representation can come from two sources: psychological insight into the nature of the cognitive process, and knowledge of the properties of various data representations for computer programs. Cognitive science brings these two sources together in the hope that a better model will come from consideration of both. In the past, psychologists have been guilty of ignoring arguments from computer science, and AI researchers have likewise ignored psychological results. SaMc is the product of a line of research that is centered around a particular, restricted form of computation – connectionist networks. The goal of this research is to provide connectionist accounts of a wide range of cognitive phenomena. Unfortunately, the desire to stay within the connectionist framework leads to models that are shaped more by the connectionist ideal than the phenomena they aim to model. Connectionism is a mechanism; when building a cognitive model we should be driven by our observations towards an appropriate mechanism, rather than trying to fit what we see into the mechanism we have chosen.

Performance

It is clear that a cognitive theory should be founded on empirical evidence, and any model based on such a theory should reproduce empirical observations. The tradition of Artificial Intelligence, however, relies much less on detailed observations of human performance than on 'intuitions' about how a problem is solved, or on some 'neat' algorithm developed for a task. In a sense, the AI model provides an 'existence proof' that there is an algorithm for the task being modeled, without making strong claims about the relationship between the real and artificial systems. A cognitive model should be more concerned with established observations of performance. The issues are what sorts of behaviour should be taken into account, and how a model's performance can be compared with that of humans.

Two experimental tasks are often used to measure reading performance. Lexical decision tasks involve showing the subjects letter strings and asking them to decide whether the letters make up a word. Naming tasks require the subject to pronounce a letter string, which again may or may not be a word. In both cases measurements of reaction times and accuracy can be made. Experimenters look at the relationship between performance on different classes of words (high/low frequency, regular/irregular letter-sound correspondence etc.) and, in developmental studies, at the change in performance over time.

Outside the psychologist's laboratory, the main measures of reading performance are the number of words that can be named correctly, and the level of understanding gained from the text. We are not concerned with understanding here (although it may play an important role in word recognition); however, naming competence can be used to evaluate a model of word recognition. We can observe two aspects of naming performance: the number, and types of words that are named consistently and the error responses made when reading. By examining these *qualitative* measures of performance in addition to the *quantitative* measures, such as reaction time experiments, we obtain a broader picture of the process of learning to read than if we limit the *types* of observations we consider.

Seidenberg and McClelland measure the 'reaction time' of their model by making an assumption about the process of naming a word:

> "We assume that overt naming involves three cascaded processes (see also Balota and Chumbley (1985)): (a) the input's phonological code is computed; (b) the computed phonological code is compiled into a set of articulatory-motor commands; (c) the articulatory-motor code is executed, resulting in the overt response. Only the first of these processes is implemented in the model. In practice, however, [the quality of] the phonological output computed by the model is closely related to observed naming latencies"

They assume that the reaction time in a naming task is related to the degree of ambiguity in the output phonological representation. This can be measured by comparing the actual output (states of the phonological units) with the desired output. Any unit that has a different activation (too high or too low) level to the desired output contributes to an error score. This error score is assumed to be proportional to the naming latency, as a highly ambiguous output will cause a delay in the construction of an articulatory-motor program.

Using the reaction time measure, Seidenberg and McClelland are able to reproduce some of the classic experiments comparing reaction times for different classes of words. Of particular interest here is the behavior of the model at different stages of development. A common observation is that the difference in processing speed for regular and irregular words changes over time. In the early stages of development irregular words take longer to pronounce than regular ones, regardless of their frequency. Later, high frequency words of all kinds have about the same naming latencies, however the regularity effect is retained for low frequency words. The performance of SaMc, measured at various times during the training regime, reproduces this observation.

In SaMc, quantitative performance is more easily observed than qualitative performance – given the timing assumptions discussed above. Determining whether an error response was made, and which word was used, is difficult due to the inherent ambiguity of the triple representation (Pinker & Prince, 1988). The phonological output produced for a given input consists of a pattern of activation over the phonological units. Each of these units represents a triple of phonetic features (Section) and can have an activation level ranging from zero to one. For a well-learned word it is likely that phonetic features not contained in the word will receive some small activation. For an unknown word, the correct phonological units may be activated, but so will many other units – perhaps less strongly. In addition to this, there is no guarantee that the phonetic triples activated can be strung together to form a pronounceable phonetic string, or that they represent one unique string (Prince & Pinker, 1988). Thus, given the representation chosen by Seidenberg and McClelland, it is difficult to examine the performance of the model in qualitative terms.

The performance of SaMc on quantitative measures is impressive. It suggests that some of the mechanisms being put forward are appropriate for reading development. However, because we are unable to examine its qualitative performance, the model cannot be fully evaluated. The reason for the lack of qualitative data is the use of a connectionist architecture and the choice of phonological representation. One interesting observation is that, when learning to read, children tend to make errors within their reading vocabulary; that is if they make an incorrect response to a word the error will be another word they can read, as opposed to a non-reading word or a random string of phonemes. This is just one observation that is difficult to account for in a connectionist model like SaMc which learns associations between orthography and phonology. If SaMc makes an error it is likely to consist of some of the correct phonemes and other spurious or wrong ones; it is highly unlikely that the response would be another complete word from the reading lexicon. As in Section , it would be possible to build a connectionist model that showed some of this behaviour. It would, however, be difficult for such a model to progress to a stage where it makes responses from outsides its reading vocabulary, as happens with most children.

Credibility

The development of word recognition skills is only a small part of the cognitive development of a child. We choose to model it separately because to model the whole would be too complex, and we assume that human cognition is separable into functional units with limited, definable interactions. This assumption may not be valid; we may be *forced* to consider the relationship between, say, word recognition and access to meaning. A single piece of work may not be able to address the wider context, but it should be consistent with models and theories of related tasks.

A credible computational model will reproduce the relevant behavior by using a mechanism that has support from other areas of research, or that is supported by substantive justification. For example, the ACT* architecture (Anderson, 1983) proposes spreading activation in a semantic network as a primary cognitive process; a number of models have been built using this architecture. Each new model not only adds to the credibility of ACT* as an architecture for cognition, but gains support for itself by using mechanisms shown to be general in their utility. Of course there are a number of competing mechanisms – ACT* and connectionism are examples – so there is no one *correct* choice.

New mechanisms and structures must be proposed if research in cognitive modelling is to progress beyond its current state. Every mechanism hides a number of assumptions about its implementation and the nature of the task being modeled. To enable evaluation of the model, these assumptions need to be made explicit; to ensure that the model is credible they need to be rigorously justified. Van Lehn (1985) puts forward competitive argumentation as a means of establishing the appropriate assumptions to be made in a model, and of providing justification for them. Competitive argumentation considers all the choices (hypotheses) available – for instance, all the ways of organising the lexicon – and selects one as the best or most general. The selection process produces a number of arguments for the chosen hypothesis which make up its justification. In this way the basic hypotheses of a theory/model are made explicit and the reasons for them are clear.

SaMc is one of a number of cognitive models implemented as connectionist networks. Other connectionist models have been constructed for learning the past tenses of English verbs (Rumelhart & McClelland, 1987), parsing natural language (Waltz & Pollack, 1985) and modelling the effects of context on letter perception (McClelland & Rumelhart, 1981). One of the goals of this research seems to be to show the general utility of connectionism as a mechanism for cognition. As such, connectionism performs creditably well on the first criteria above; the suggested relationship of SaMc to the rest of cognition is clear: it is a connectionist architecture, using similar mechanisms to connectionist models of other skills. Seidenberg and McClelland fail to explicate and provide justification for the assumptions upon which their model is based; assumptions about the interfaces to the model and about how learning takes place in the environment in which the model exists.

SaMc requires a particular form of input and produces a particular output, it is not clear how or where these could be generated. Implicit in the need for letter triple inputs is an assumption that word recognition is letter mediated – there is a process (part of visual perception?) that constructs a letter based description from a pattern of visual stimuli. The letter based representation is also required to be complete (contains all and only the letters seen) and positionally stable (the same word will always produce the same representation on different presentations)[5]. Each of these characteristics of the input representation can be argued for and against; Seidenberg

[5]In fact there is some *noise* in the input representation – it is not always complete and positionally stable – the motivation for this noise is to enable the learning algorithm to make the appropriate generalisations. The characteristics of the noise are not based on any psychological considerations.

and McClelland may claim that they are all clearly acceptable assumptions. Our criticism is that the important assumptions about the characteristics of the input representation are not made explicit and are given no justification.

Claiming that the credibility of a model (as defined here) is important makes a methodological point about how to do research in cognitive science. In order for the details of the hypotheses on which a model is based to be made clear, care must be taken when formulating the model. We believe that by doing this, better models will result and more will be learned from individual pieces of research.

CONCLUSION

Computational models of learning provide a valuable contribution to cognitive science: they require that most of the assumptions of a theory be made explicit. A computer model can be used to predict the behavior of a complex system operating in a complex environment. There is a tradition of building computational models of cognitive tasks in the field of Artificial Intelligence; it is tempting to think that, once built, such a program constitutes a theory of the cognitive task. If we are to use computational models in cognitive science we must be aware of the role they play, and how they can be used to further a theory of cognition.

Seidenberg and McClelland present a model of word recognition that learns to associate orthographic and phonological representations of words. The model is implemented as a connectionist network. The results of experiments carried out using the model suggest that it has captured some important aspects of the word recognition task; its behavior is very close to that of a human reader on a number of experimental tasks. We have used this model as a starting point in developing a set of criteria by which a cognitive developmental model can be evaluated. These criteria, and our evaluation of SaMc, are summarised here.

Firstly, the environment in which the model learns should approximate that of a child. Seidenberg and McClelland do not attempt to capture some important features of a child's reading vocabulary; the words used to train the model are chosen to restrict the complexity of the spelling to sound correspondences which the model has to learn. Also, children's environments change as they develop, this change is not accounted for or taken advantage of in SaMc. It appears that the shape of the environment presented to the model was determined more by constraints imposed by the connectionist implementation, than by the real environment in which a child learns.

The representation of the lexicon used in the model needs to be sufficiently rich to capture all the stages of learning; from the early 'whole word' recognition stage to the mature stage, recognising words by analogy. SaMc uses a connectionist network which is able to adapt as it learns to capture generalities in the patterns it is exposed to. However, there are milestones on the road to mature word recognition that SaMc does not, and cannot, pass because of the nature of learning in current connectionist networks.

When evaluating a cognitive model, it is important to compare its performance with that of the cognitive systems it claims to model. Too often researchers concentrate on reproducing quantitative behaviour – timing measurements for various experimental tasks – and ignore the equally important qualitative data. In fact, qualitative data is arguably *more* important – the right timings for the wrong (qualitative) answers is no validation at all. Seidenberg and McClelland take great pains to reproduce a number of important findings and also make predictions which are backed up by subsequent experimental work. Their model is incapable of making statements about qualitative performance – such as what types of errors are made in response to unknown words. The model is designed to reproduce timing data, and does so admirably, but we are missing an important part of the evaluation by not being able to examine qualitative performance.

Finally, we must be able to believe in a model. There should be no hidden assumptions waiting to demolish its credibility. Assumptions need to be made when building a computational model. It is important to realise that some of these assumptions are properly part of the theory, others are merely conveniences. The latter should not be such that changing them produces drastic changes in the behaviour and predictions of the model. SaMc embodies many assumptions which are not properly justified; this does not mean that they could not be justified. In the case of Seidenberg and McClelland, it is hard to determine which assumptions are theoretical claims and which arise from the implementation requirements of SaMc.

In summary, SaMc is a successful model in many respects, but the ideas expressed within it are held back by the connectionist implementation. This introduces constraints where none belong: the constraints on the implementation should come from the world being modeled rather than the mechanism chosen for implementation.

ACKNOWLEDGEMENTS

I would like to thank Peter Andreae and Michael Moses of the Department of Computer Science and Brian Thompson of the Department of Education, Victoria University of Wellington, for their comments on the form and content of this paper. I am grateful to Dr. G.J. Dalenoort, Prof. P van Geert and Dr. Z. Schreter for comments on an earlier draft of this paper.

REFERENCES

Anderson, J. R. (1983). *The Architecture of Cognition.* Harvard University Press, Cambridge, MA.

Balota, D. A. & Chumbley, J. I. (1985). The locus of the word frequency effect in the pronunciation task: lexical access and/or production. *Journal of Memory and Language*, *24*, 89–106.

Campbell, R. (1987). Oral reading errors of two young beginning readers. *Journal of Research in Reading*, *10*(2), 144–155.

Elman, J. (1988). Finding structure in time. CRL Technical Report 8801, Center for Research in Language, University of California, San Diego.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In Patterson, K., Marshall, J. C., & Coltheart, M. (Eds.), *Surface Dyslexia and Surface Dysgraphia in Adults and Children.* Lawrence Earlbaum.

Glushko, R. J. (1979). The organisation and activation of orthographic knowledge in reading aloud. *J. Experimental Psychology: Human Perception and Performance*, *5*, 674–691.

Hinton, G. E. (1987). Connectionist learning procedures. Technical Report CMU-CS-87-115, Computer Science Department, Carnegie Mellon University.

Humphreys, G. W. & Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *The Behavioral and Brain Sciences*, *8*, 689–740.

Kucera, H. & Francis, N. (1967). *Computational analysis of present-day American English.* Brown University Press.

Marsh, G., Morton, F., Welch, V., & Desberg, P. (1980). The development of strategies in spelling. In Frith, U. (Ed.), *Cognitive Processes in Spelling*, chap. 15, pp. 339–354. Academic Press.

McClelland, J. L. & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: part 1. an account of basic findings. *Psychological Review*, *88*(5), 375–407.

Pinker, S. & Prince, A. (1988). On language and connectionism: analysis of a parallel distributed processing model of language aquisition. *Cognition*, *28*, 73–194.

Prince, A. & Pinker, S. (1988). Wickelphone ambiguity. *Cognition*, *30*, 189–190.

Rumelhart, D. E. & McClelland, J. L. (1987). Learning the past tenses of English verbs: implicit rules of parallel distributed processing?. In MacWhinney, B. (Ed.), *Mechanisms of Language Aquisition*. Lawrence Earlbaum.

Seidenberg, M. & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568.

Sejnowski, T. J. & Rosenberg, C. R. (1986). NETTALK: a parallel network that learns to read aloud. Technical Report JHU/EECS-86/01, Johns Hopkins University EE and CS.

Seymour, P. & Elder, L. (1986). Beginning reading without phonology. *Cognitive Neuropsychology*, *3*(1), 1–36.

Seymour, P. (1984). Reading and spelling development. In *Developmental Dyslexia: A Cognitive Experimental Analysis*. Venice Meeting on Cognitive Neuropsychology of Language.

Thompson, G. B. (1982). Initial vocabularies for reading.. Set, Research Information for Teachers, number 2, item 3, Australian Council for Educational Research.

Van Lehn, K. (1985). *Felicity conditions for Human skill aquisition: Validating an AI based theory*. Ph.D. thesis, MIT.

Waltz, D. & Pollack, J. (1985). Massively parallel parsing: a strongly interactive model of natural language interpretation. *Cognitive Science*, *9*, 51–74.