What do connectionnist simulations tell us?

Gobet, F. (1999). What do connectionnist simulations tell us? *Journal of Child Language, 26*, 242-245.

Fernand Gobet

Department of Psychology

University of Nottingham

Nottingham NG7 2RD

England

(0115) 951 5402

frg@psyc.nott.ac.uk

What do connectionnist simulations tell us?

In his review, Rispoli's main concern is that Elman et al.'s book will aggravate the degree of polarisation in developmental psycholinguistics. I cannot really comment on this worry, as developmental psycholinguistics is not my field. Instead, I will discuss some questions more related to my background--the role of computational modelling in Elman et al.'s approach.

Elman et al.'s ambitious goal is to propose theories of cognitive development that are grounded in our knowledge of biology. This if of course what the great Jean Piaget tried to achieve during his lifetime—unsuccessfully, as we know. Elman et al.'s advantage over Piaget is that they have a set of computational tools, connectionnism, which both allows them to specify theories precisely and to study complex behaviours (such as epigenesis, where innate and environmental factors interact to create new levels of complexity) that are just beyond the (unaided) human mind. Even though I will highlight some of the weaknesses of their approach below, I should emphasise that reading their book was an exciting and enjoyable experience.

As noted by Rispoli, there are important problems with the simulations reported by Elman et al. Rispoli focuses on simulations of past tense acquisition and syntax acquisition, but the problems are by no means limited to these areas. I will briefly consider two recent developments in neural net research, one taken from the field of language acquisition, and one from elsewhere, which underscore some of the difficulties of the simulations discussed in the book.

The influential simulations reported first in Elman (1993), stressing the "importance of starting small", play a significant role in the book, because they support one of its key ideas, that time-related constraints on perceptual capacity and memory capacity may facilitate the acquisition of language (or of other complex skills) in that they limit the search space faced by biological brains and by artificial neural networks. Elman found that a recurrent network could not acquire a simple, artificial grammar containing relative clauses when full sentences were given as input, but that it could do so either when the complexity of the input was increased gradually, short sentences being presented first, or when the temporal window with which the network could process information was limited to three or four words at the beginning and then progressively extended to its full size. However, Rohde and Plaut (1997), using essentially the same type of recurrent network as Elman, were unable to replicate this finding. They found that their network could learn better when it was presented with full sentences than when presented with an incremental regime. They concluded that, if anything, delaying complex information was often an hindrance to the network.

The second example is taken from simulations of the balance beam task (a classic task in the developmental literature) first reported by McClelland and Jenkins (1991). These simulation occupy an important place in the book, because they show that a neural net using a continuous, non-linear learning function can display behaviour that looks stage-like and discontinuous. Unfortunately for this claim, however, Raijmakers, van Koten, and Molenaar (1996) have recently shown that the stage-wise behaviour of the network is an artefact of the measures used by McClelland and Jenkins. When a proper set of measures is employed, with strict criteria for assessing discontinuity (criteria based on the mathematical catastrophe theory) it was shown that the neural net displayed no sign of stage-like behaviour.

In general, the simulations in Elman et al. come out best when they explore the dynamics of neural networks for themselves, without reference to empirical data. When they try to simulate data, they are plagued by the many problems noted by Rispoli. In this respect, it is unfortunate that Elman et al. do not discuss alternative computational approaches, in particular adaptive production systems (e.g., Langley, 1987), which are, like neural nets, self-organising systems displaying non-linear behaviour. In doing so, they could have contrasted both the strengths and weaknesses of neural nets, and perhaps found leads for further research. An obvious domain where this ecumenical approach would have helped is high-level cognition, such as reasoning, planning, or scientific thinking, about which connectionism is mostly silent, as noted by Elman et al. themselves (p. 394).

It is somewhat disappointing that Rispoli does not discuss the emphasis given by the book to recent advances in developmental neurobiology, because these constitute a cornerstone of Elman et al.'s argument. I believe that Elman et al. are right to challenge what they call representational innateness, which plays a prominent role in developmental psychology in general (e.g. the work of Carey and Spelke), and in developmental psycholinguistics in particular (e.g. the work of Chomsky and Pinker). Their detailed criticism of the data supposed to support innate representations (cf. Chapter 7, where the matter is discussed with respect to language) and the evidence they adduce in favour of epigenesis (plasticity and equipotentiality of the brain, as well as the fact that the size of the human genome is not large enough to encode representational information) is compelling, although these questions are of course currently highly controversial in biology and in neuropsychology. Even so, their criticisms pose a welcome challenge to theoreticians who use "explanations" based on innate factors, but do not provide detailed mechanisms.

Elman et al.'s argument becomes much weaker when they claim that connectionism offers a satisfactory explanation for epigenesis, however. Most of the models discussed in the book abstract too much from biological complexity to make this claim plausible. This is the case even in Chapter 6, which attempts to focus specifically on biologically plausible models. In particular, the simulations on the importance of starting small (Elman, 1993) are discussed at length in this chapter, as they are used to illustrate the importance of chronotopic constraints in the development of natural and artificial networks. As noted above, however, Rohde and Plaut (1997) were unable to replicate Elman's findings, showing that starting with full complexity leads to better results.

Elman et al. offer a nice introduction to non-linear dynamics, to neural nets, and to recent developments in developmental neurobiology. The attempt to put these three fields together into a single research framework does not come out convincingly, however. I think that this failure is due both to the complexity of these fields taken individually, and to the fact that our knowledge of the biological basis of development does not (yet) offer the constraints hoped for. Given that the simulations reported in the book are sometimes not replicable and often do not account for the data satisfactorily--which both are of course obvious requirements for a model or theory that calls itself successful--I have to agree with Rispoli that Elman et al.'s book must be taken more as a position statement than as a scientific breakthrough.

References

Elman, J. L. (1993). Learning and development in neural networks: the importance of starting small. *Cognition*, 48, 71-99.

Langley, P. (1987). A general theory of discrimination learning. In D. Klahr, P. Langley & R. Neches (Eds.), *Production system models of learning and development*. Cambridge, MA: The MIT Press.

McClelland, J. L., & Jenkins, E. (1991). Nature, Nurture and Connections: Implications of connectionist models for cognitive development. In K. VanLehn, (Ed.).*Architectures for intelligence*. Hillsdale, NJ: Erlbaum.

Raijmakers, M., van Koten, S., & Molenaar, P. (1996). On the validity of simulating stagewise development by means of PDP networks: Application of catastrophe analysis and an experimental test of rule-like network performance. *Cognitive Science, 20,* 101-136.

Rohde, D. L. T. and D. C. Plaut (1997). Simple recurrent networks and natural language: How important is starting small? *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 656-661). University of Stanford. Mahwah, NJ: Erlbaum.