

CENTER FOR RESEARCH IN LANGUAGE

February 1996

Vol. 10, No. 4

The Newsletter of the Center for Research in Language, University of California, San Diego, La Jolla CA 92093
Tel: (619) 534-2536 • E-mail: crl@crl.ucsd.edu • WWW: <http://crl.ucsd.edu/newsletter.html>

• • •

FEATURE ARTICLE

Development in a Connectionist Framework: Rethinking the Nature-Nurture Debate

Kim Plunkett

Oxford University

ANNOUNCEMENTS

Postdoctoral Fellowship: The Center for Research in Language announces postdoctoral fellowships in Language, Communication, and Brain. The fellowships are supported by the National Institutes of Health and provide an annual stipend ranging from \$19,608 to \$32,300. Deadline for submission of applications is May 3, 1996.

WWW: The CRL Newsletter is now on World Wide Web at URL: <http://crl.ucsd.edu/newsletter.html>

EDITOR'S NOTE

This newsletter is produced and distributed by the **CENTER FOR RESEARCH IN LANGUAGE**, a research center at the University of California, San Diego that unites the efforts of fields such as Cognitive Science, Linguistics, Psychology, Computer Science, Sociology, and Philosophy, all who share an interest in language. We feature papers related to language and cognition (1-10 pages, distributed via e-mail and the World Wide Web) and welcome response from friends and colleagues at UCSD as well as other institutions.

SUBSCRIPTION INFORMATION

If you are currently receiving a hardcopy of the newsletter and have access to e-mail, please help us save printing and mailing costs by forwarding your e-mail address to CRL. If you require a hardcopy in addition, please request it and we will be happy to send you one.

If you know of others who would be interested in receiving the newsletter, please forward the e-mail or postal mailing address. Please forward correspondence to:

Jay Moody, Editor
Center for Research in Language, 0526
9500 Gilman Drive, University of California, San Diego 92093-0526
Telephone: (619) 534-2536 • E-mail: crl@crl.ucsd.edu

Back issues of this newsletter are available from CRL in hard copy as well as soft copy form. Papers featured in previous issues include the following:

Zai and Ba Constructions in Child Mandarin

Ping Li

Center for Research in Language, UCSD
Vol. 5, No. 5, July 1991

Hitting the Right Pitch: A Meta Analysis of Sentence Context on Lexical Access

Mark St. John

Department of Cognitive Science, UCSD
Vol. 5, No. 6, August 1991

What is Who Violating? A Reconsideration of Linguistic Violations in Light of Event Related Potentials

Marta Kutas

Cognitive Science and Neurosciences, UCSD

Robert Kluender

Department of Linguistics, UCSD
Vol. 6, No. 1, October 1991

Learning to Recognize and Produce Words: Towards a Connectionist Model

Michael Gasser

Department of Computer Science, Indiana University
Vol. 6, No. 2, November 1991

PDP Learnability and Innate Knowledge of Language

David Kirsh

Department of Cognitive Science, UCSD
Vol. 6, No. 3, December 1991

Why It Might Pay to Assume That Languages Are Infinite

Walter J. Savitch

Department of Computer Science, UCSD
Vol. 6, No. 4, April 1992

Mental Spaces and Constructional Meaning

Claudia Brugman

Center for Research in Language, UCSD
Vol. 7, No. 1, October 1992

Is Incest Best? The Role of Pragmatic Scales and Cultural Models in Abortion Rhetoric

Seana Coulson

Department of Cognitive Science, UCSD
Vol. 7, No. 2, January 1993

Marking Oppositions in Verbal and Nominal Collectives

Suzanne Kemmer

Department of Linguistics, UCSD
Vol. 7, No. 3, September 1993

Connectionist Representations: The State of the Art

Daniel Memmi

LIMSI-CNRS
Orsay (France)
Vol. 8, No. 1, November 1993

Abstract Better Than Concrete: Implications for the Psychological and Neural Representation of Concrete Concepts

Sarah D. Breedin, Eleanor M. Saffran, and H. Branch Coslett

Center for Cognitive Neuroscience, Department of Neurology, Temple University
Vol. 8, No. 2, April 1994

Analogic and Metaphoric Mapping in Blended Spaces: Menendez Brothers Virus

Seana Coulson

Department of Cognitive Science, UCSD
Vol. 9, No. 1, February 1995

In Search of the Statistical Brain

Javier Movellan

Department of Cognitive Science, UCSD
Vol. 9, No. 2, March 1995

Connectionist Modeling of the Fast Mapping Phenomenon

Jeanne Milostan

Computer Science and Engineering, UCSD
Vol. 9, No. 3, July 1995

Representing the Structure of a Simple Context-Free Language in a Recurrent Neural Network: A Dynamical Systems Approach

Paul Rodriguez

Department of Cognitive Science, UCSD
Vol. 10, No. 1, October 1995

A Brain Potential Whose Latency Indexes the Length and Frequency of Words

Jonathan W. King

Cognitive Science, UCSD

Marta Kutas

Cognitive Science and Neurosciences, UCSD
Vol. 10, No. 2, November 1995

Bilingual Memory: A Re-Revised Version of the Hierarchical Model of Bilingual Memory

Roberto R. Heredia

Center for Research in Language, La Jolla, CA
Vol. 10, No. 3, January 1996

Development in a Connectionist Framework: Rethinking the Nature-Nurture Debate

Kim Plunkett
Oxford University

A Developmental Paradox

Two findings in developmental psychology stand in apparent conflict. Piaget (1952) has shown that at a certain stage in development, children will cease in their attempts to reach for an object when it is partially or fully covered by an occluder. This finding is observed in children up to the age of about 6 months and is interpreted to indicate that the object concept is not well-established in early infancy. The object representations that are necessary to motivate reaching and grasping behavior are absent. In contrast, other studies have shown that young infants will express surprise when a stimulus array is transformed in such a way that the resulting array does not conform to reasonable expectations. For example, change in heart rate, sucking or GSR, is observed when an object, previously visible, fails to block the path of a moving drawbridge or a locomotive fails to reappear from a tunnel, or has changed colour when it reappears (Baillargeon, 1993; Spelke, 1994). These results are interpreted as indicating that important representations of object properties such as form, shape and the capacity to block the movement of other objects are already in place by 4 months of age. The conflict in these findings can be stated as follows: Why should the infant cease to reach for a partially or fully concealed object when it already controls representational characteristics of objects that confirm the stability of object properties over time, and that predict the interaction of those represented properties with objects that are visible in the perceptual array?

One answer to this conflict is that Piaget grossly underestimated young children's ability to retrieve hidden objects. However, this answer is no resolution to the conflict: Piaget's findings are robust. Alternatively, one might question Piaget's interpretation of his results. Young infants know a lot about the permanent properties of objects but recruiting object representations in the service of a reaching task requires additional sensorimotor skills which have little to do with the infant's understanding of the permanence of objects. Again, this response must be rejected. Young infants who are in full command of the skill to reach and grasp a visible object still fail to retrieve an object which is partially or fully concealed (von Hofsten, 1989). Motor skills are not the culprit here. The capacity to relate object knowledge to other domains seems to

be an important part of object knowledge itself. Object knowledge has to be accessed and exercised.

A Resolution

A resolution of the conflict can be found in considering some fundamental differences in the nature of the two types of task that infants are required to perform. In experiments that measure "surprise" reactions to unusual object transformations such as failure to reappear from behind an occluder, the infant is treated as a passive observer (Baillargeon, 1993). In essence, the infant is evaluated for its expectations concerning the future state of a stimulus array. Failure of expectation elicits surprise. In the Piagetian task, the infant is required to actively transform the stimulus array. To achieve this, not only must the infant know where the object is but she must be able to coordinate that information with knowledge about the object's identity—typically, the infant reaches for objects she wants. We suppose that this coordination is relatively easy for visible objects, because actions are supported by externally available cues. However, when the object is out of sight, the child has to rely on internal representations of the object's identity and position. We assume that the internal representations for object position and identity develop separately. This assumption is motivated by recent neurological evidence that spatial and featural information is processed in separate channels in the human brain—the so-called 'what' and 'where' channels (Ungerlieder & Mishkin 1982). In principle, the child could demonstrate knowledge of an object's position without demonstrating knowledge about its identity, or *vice versa*. Surprise reactions might be triggered by failure of infant expectations within either of these domains. For example, an object may suddenly change its featural properties or fail to appear in a predicted position. Internal representations are particularly important when the object is out of sight. Hence, we might expect infants to have greater difficulty performing tasks that involve the coordination of spatial and featural representations—such as reaching for hidden objects—when these representations are only partially developed.

Building a model

The resolution outlined in the previous section constitutes a theory about the origins of infants' surprise reactions to

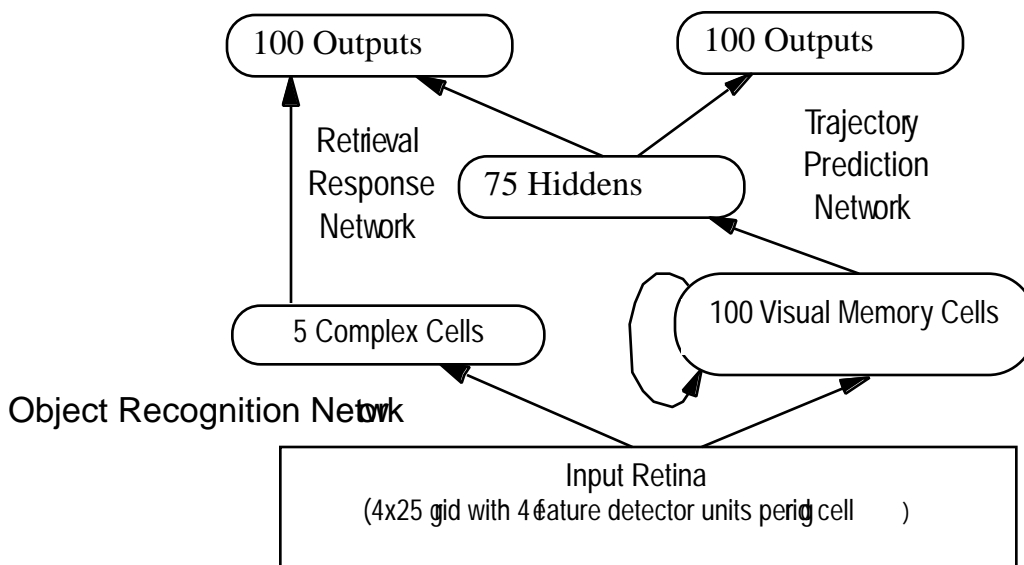


Figure 1: The modular neural network (Mareschal et al., 1995) used to track and initiate reaching responses for visible and hidden objects. An object recognition network and a visual tracking network process information from an input retina. The object recognition network learns spatially invariant representations of the objects that move around the retina. The visual tracking network learns to predict the next position of the object on the retina. The retrieval response network learns to integrate information from the other two modules in order to initiate a reaching response. The complete system succeeds in tracking visible objects before it can predict the reappearance of hidden objects. It also succeeds in initiating a reaching response for visible objects before it learns to reach for hidden objects.

objects' properties (spatial or featural) which do not conform to expectations and attempts to explain why these surprise reactions precede the ability to reach for hidden objects even though they possess the motor skills to do so. Mareschal, Plunkett & Harris (1995) have constructed a computational model that implements the ideas outlined in this theory. The model consists of a complex neural network that processes a visual image of an object that can move across a flat plane. Different types of objects distinguished by a small number of features appear on the plane one at a time. These objects may or may not disappear behind an occluder. All objects move with a constant velocity so that if one disappears behind an occluder, it will eventually reappear on the other side. Object velocities can vary from one presentation to the next.

The network is given two tasks. First, it must learn to predict the next position of the moving object, including its position when hidden behind an occluder. Second, the network must learn to initiate a motor response to reach for an object, both when visible and when hidden. The network is endowed with several information processing capacities that enable it to fulfil these tasks. The image of the object moving across the plane is processed by two separate modules. One module learns to form a spatially invariant representation of the object so that it can recognise its identity irrespective of its position on the plane (Foldiak 1991). The second module learns to keep track of the object but loses all information about the object's identity (Ungerlieder & Mishkin 1982). This second module

does all the work that is required to predict the position of the moving object. However, in order to reach for an object, the network needs to integrate information about the object's identity and its position. Both modules are required for this task. Therefore, the ability to reach can be impeded either because the representations of identity and position are not sufficiently developed or because the network has not yet managed to properly integrate these representations in the service of reaching.

Given the additional task demands imposed on the network for reaching it would seem relatively unsurprising to discover that the network learns to track objects before it learns to reach for them. The crucial test of the model is whether it is able to make the correct predictions about the late onset of reaching for hidden objects relative to visible objects. In fact, the model makes the right predictions for the order of mastery in tracking and reaching for visible and hidden objects. It quickly learns to track and reach for visible objects, tracking being slightly more precocious than retrieval. Next, the network learns to track occluded objects as its internal representations of position are strengthened and it is able to "keep track" of the object in the absence of perceptual input. However, the ability to track hidden objects together with the already mastered ability to reach for visible objects does not guarantee mastery of reaching for hidden objects. The internal representations that control the integration of spatial and

featural information require further development before this ability is mastered.

Evaluating the Model

Notice how this modelling endeavour provides a working implementation of a set of principles that constitute a theory about how infants learn to track and reach for visible and hidden objects. It identifies a set of tasks that the model must perform and the information processing capacities required to perform those tasks. All these constitute a set of assumptions that are not explained by the model. However, given these assumptions, the model is able to make correct predictions about the order of mastery of the different tasks. The model implements a coherent and accurate (not necessarily true—the assumptions might be wrong) theory. However, this model just like any other has a number of free parameters which the modeller may ‘tweak’ in order to achieve the appropriate predictions. It is necessary to derive some novel predictions which can be tested against new experimental work with infants, in order to evaluate the generality of the solution the model has found. This model makes several interesting predictions including improved tracking skills at higher velocities and imperviousness to unexpected feature changes while tracking. The first experimental prediction has been confirmed (see Mareschal, Harris & Plunkett 1995) while the second prediction is currently being tested. This instance of model building and evaluation thus seems to support the initial insight that children’s object representations develop in a fragmentary fashion, and that the development of these fragments of knowledge shape infant performance on various tasks in line with their manner of involvement in the tasks concerned.

Connectionist Insights

The model described in the previous section is an example of a computer simulation that uses the learning capabilities of artificial neural networks to construct internal representations of a training environment in the service of several tasks (reaching and tracking). Neural networks are particularly good at extracting the statistical regularities of a training environment and exploiting them in a structured manner to achieve some goal. They consist of a well-specified architecture driven by a learning algorithm. The connections or weights between the simple processing units that make up the network are gradually adapted over time in response to localised messages from the learning algorithm. The final configuration of weights in the network constitutes what it knows about the environment and the tasks it is required to perform.

Connectionist modelling provides a flexible approach to evaluating alternative hypotheses concerning the start state of the organism (or what we may think of as its innate endowment), the effective learning environment that the

organism occupies and the nature of the learning procedure for transforming the organism into its mature state. The start state of the organism is modelled by the choice of network architecture and computational properties of the units in the network. There are a wide range of possibilities that the developmentalist can choose between. The effective learning environment is determined by the manner in which the modeller chooses to define the task for the network. For example, the modeller must decide upon a representational format for the pattern of inputs and outputs for the network, and highlight the manner in which the network samples patterns from the environment. These decisions constitute precise hypotheses about the nature of the learning environment. Finally, the modeller must decide how the network will learn. Again, a wide variety of learning algorithms are available to drive weight adaptation in networks. Any particular connectionist model embodies a set of decisions governing all of these factors which are crucial for specifying clearly one’s theory of development. Quite small changes in one of the choices can have dramatic changes for the performance of the model—some of them quite unexpected. Connectionist modelling offers a rich space for exploring a wide range of developmental hypotheses.

In the remainder of this article I will briefly review some connectionist modelling work that has explored some important areas in the hypothesis space of developmental theories. I aim to underscore four main lessons or insights that these models have provided:

1. When constructing theories in psychology, we use behavioural data from experiments or naturalistic observation as the objects that our explanations must fit. We attempt to infer underlying mechanisms from overt behaviour. Connectionist modelling encourages us to be suspicious of the explanations we propose. Often, networks surprise us with the simplicity of the solution they discover to apparently complex tasks—sometimes, leading us to the conclusion that learning may not be as difficult as we thought.
2. When we see new forms of behaviour emerging in development, we are tempted to conclude that some radical change has occurred in the mechanisms governing that behaviour. Connectionist modelling has shown us that small and gradual internal changes in an organism can lead to dramatic non-linearities in its overt behaviour—new behaviour need not mean new mechanisms.
3. Theories of development are often domain specific. Behaviours that are discrete and associated with distinguishable modalities, promote explanations that do not reach beyond the specifics of those modalities or domains. These encapsulated accounts often emphasise the impoverished character of the learning environment and lead to complex specifications of the organism’s start state. Connectionist models provide a framework for investigating

the interaction between modalities and a formalism for entertaining distributed as well as domain specific accounts of developmental change. This approach fosters an appreciation of developing systems in which domain specific representations emerge from a complex interaction of the organism's domain-general learning capacities with a rich learning environment.

4. Complex problems seem to require complex solutions. Mastery of higher cognitive processes appears to require the application of complex learning devices from the very start of development. Connectionist modelling has shown us that placing limitations on the processing capacity of developing systems during early learning can actually enhance their long-term potential. The ignorance and apparent inadequacies of the immature organism may, in fact, be highly beneficial for learning the solutions to complex problems. Small is beautiful.

Inferring Mechanisms from Behaviour

Children make mistakes. Developmentalists use these mistakes as clues to discover the nature of the mechanisms that drive correct performance. For example, in learning the past tense forms of irregular verbs or plurals of irregular nouns, English children may sometimes overgeneralise the “-ed” or “s” suffixes to produce incorrect forms like “hitted” or “mans”. These errors often occur after the child has already produced the irregular forms correctly, yielding the well-known U-shaped profile of development.

A Dual-Mechanism Account

A natural interpretation of this pattern of performance is to suggest that early in development, the child learns irregular forms by rote, simply storing in memory the forms that she hears in the adult language. At a later stage, the child recognises the regularities inherent in the inflectional system of English and re-organises her representation of the past tense or plural system to include a qualitatively new device that does the work of adding a suffix, obviating the need to memorise new forms. During this stage, some of the original irregular forms may get sucked into this new system and suffer inappropriate generalisation of the regular suffix. Finally, the child must sort out which forms cannot be generated with the new rule-based device. They do this by strengthening their memories for the irregular forms which can thereby block the application of the regular rule and eliminate overgeneralisation errors (Pinker & Prince 1988).

This account of the representation and development of past tense and plural inflections in English assumes that two qualitatively different types of mechanism are needed to capture the profile of development in young children—a rote memory system to deal with the irregular forms and a symbolic rule system to deal with the rest. The behavioural dissociation between regular and irregular forms—children make mistakes on irregular forms but not on regular forms—make the idea of two separate mechanisms very appealing. Double dissociations between regular and irregular forms in disordered populations add to the strength of the claim that separate mechanisms are responsible for different types of

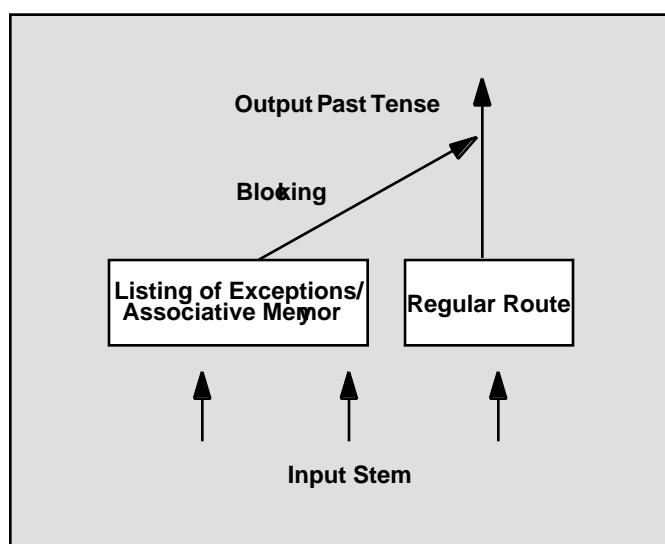


Figure 2: The dual-route model for the English past tense (Pinker & Prince 1988). The model involves a symbolic regular route that is insensitive to the phonological form of the stem and a route for exceptions that is capable of blocking the output from the regular route. Failure to block the regular route produces the correct output for regular verbs but results in overgeneralisation errors for irregular verbs. Children must strengthen their representation of irregular past tense forms to promote correct blocking of the regular route.

errors: in some language disorders children may preserve performance on irregular verbs but not on regulars while in other disorders the opposite pattern is observed.

Although the evidence is consistent with the view that a dual-route mechanism underlies children's acquisition of English inflectional morphology, this is no proof that the theory is correct. There may be other types of mechanistic explanations for these patterns of behaviour and development. Connectionist modelling offers a tool for exploring alternative developmental hypotheses.

Single-mechanism account

One of the earliest demonstrations of the learning abilities of neural networks was for English past tense acquisition. Rumelhart & McClelland (1986) suggested that the source of children's errors in learning past tense forms was to be found in their attempts to systematise the underlying relationship that holds between the verb's stem and its past tense form. For most verbs in English, the sound of the stem does not affect the past tense form. You just add "ed" on the end. However, there is a small subset of verbs which exhibit a different relationship between stem and past tense form. For example, there is a set of no change verbs where the stem and past tense forms are identical (*hit*→*hit*). All these verbs end in an alveolar consonant (/t/ or /d/). Other verbs undergo a particular type of vowel change (*ring*→*rang*, *sing*→*sang*), apparently triggered by the presence of the rhyme *-ing* in the stem. Neural networks are particularly good at picking up on

these types of regularities, so Rumelhart & McClelland trained a simple network to produce the past tense forms of verbs when presented with their stems. The details of the learning procedure and network architecture are not important here (see Plunkett 1995 for a detailed review of this and related models).

What is important is to note that Rumelhart & McClelland were successful in training the network to perform the task and that *en route* to learning the correct past tense forms of English verbs, the network made mistakes that are similar to the kind of mistakes that children make during the acquisition of inflectional morphology. Furthermore, the network did not partition itself into qualitatively distinct devices during the process of learning—one for regular verbs and one for irregular verbs. The representation of both verb types seemed to be distributed throughout the entire matrix of connections in the network. Nevertheless, a behavioural dissociation between regular and irregular verbs was observed in the network. Most of its errors occurred on irregular verbs.

More recently, Marchman (1993) has shown that damage to a network trained on the past tense problem results in further dissociations between regular and irregular forms: production of irregular forms remains intact while production of regular verbs deteriorates, mimicking patterns of performance observed in disordered populations. As with the Rumelhart & McClelland model, the representation of regular and irregular verbs was distributed throughout the network, i.e., there was

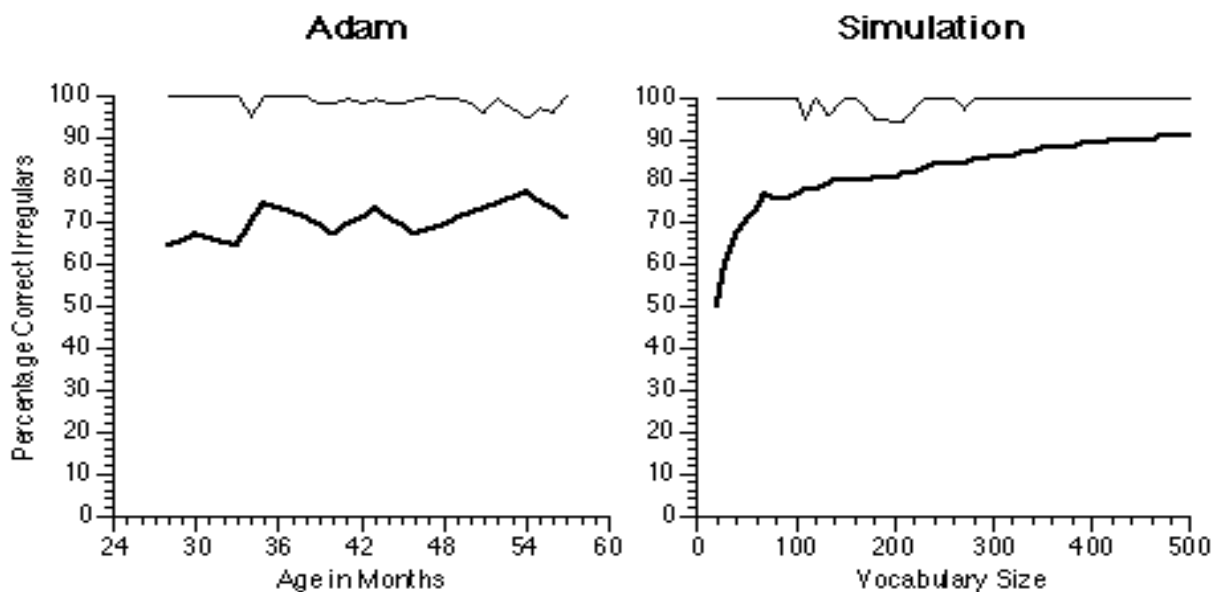


Figure 3: Network overregularization errors on irregular verbs as found in the Plunkett & Marchman (1993) simulation compared to those produced by one of 83 children analysed by Marcus, Ullman, Pinker, Hollander, Rosen & Xu (1992). The thick line indicates the percentage of regular verbs in the child's/network's vocabulary at various points in learning. Note the initial period of error free performance and overall low error rate characteristic of the developmental profiles for the model and child. Plunkett and Marchman (1993) also demonstrated that the *types* of errors that occurred in the model closely resembled the types of errors produced by the children studied by Marcus et al., (1992).

no evidence of dissociable mechanisms.

As it turns out, there were a lot of fundamental design problems with the Rumelhart & McClelland model that made it untenable as a realistic model of children's acquisition of the English past tense (Pinker & Prince 1988). Some of these problems have been fixed, some haven't (MacWhinney & Leinbach 1991, Plunkett & Marchman 1991, 1993, Cottrell & Plunkett 1994). However, the basic insight that the original model offered still remains: The observation of behavioural dissociations in some domain of performance does not necessarily imply the existence of dissociable mechanisms driving those dissociations in behaviour. Behavioural dissociations can emerge as the result of subtle differences in the graded representations constructed by these networks for different types of tasks.

Of course, just because one can train a network to mimic children's performance in learning the past tense of English verbs, does not mean that children learn them the same way as the network. The relatively simple learning system that Rumelhart & McClelland and other researchers have used to model children's learning may underestimate the complexity of the resources that children bring to bare on this problem. However, the neural network model does show that, in principle, children could use a relatively simple learning system to solve this problem. The modelling work has thereby enriched our understanding of the range and types of mechanism that might drive development in this domain.

Discontinuities in Development

Developmentalists often interpret discontinuities in behaviour as manifesting the onset of a new stage or phase of development (Piaget 1955; Karmiloff-Smith 1979; Siegler 1981). The child's transition to a new stage of development is usually construed as the onset of a new mode of operation of the cognitive system, perhaps as the result of the maturation of some cognitively relevant neural sub-system. For example, the vocabulary spurt that often occurs towards the end of the child's second year has been explained as the result of an insight (McShane 1979), in which the child discovers that objects have names. Early in development, the child lacks the necessary conceptual machinery to link object names with their referents. The insight is triggered by a switch that turns on the naming machine. Similar arguments have been offered to explain the developmental stages through which children pass in mastering the object concept, understanding quantity and logical relations.

It is a reasonable supposition that new behaviours are caused by new events in the child, just as it is reasonable to hypothesise that dissociable behaviours imply dissociable mechanisms. However, connectionism teaches us that new behaviours can emerge as a result of gradual changes in a simple learning device. It is well known that the behaviour of

dynamical systems unfolds in a non-linear and unpredictable fashion (van Geert 1991). Neural networks are themselves dynamical systems and they exhibit just these non-linear properties.

Plunkett, Sinha, Møller & Strandsby (1992) trained a neural network to associate object labels with distinguishable images. The images formed natural (though overlapping) categories so that images that looked similar tended to have similar labels. The network was constructed so that it was possible to interrogate it about the name of an object when only given its image (call this production) or the type of image when only given its name (call this comprehension). Network performance during training resembled children's vocabulary development during their second year. During the early stages of training, the network was unable to produce the correct names for most objects—it got a few right but improvement was slow. However, with no apparent warning, production of correct names suddenly increased until all the objects in the network's training environment were correctly labelled. In other words, the network went through a vocabulary spurt. The network showed a similar improvement of performance for comprehension, except that the vocabulary spurt for comprehension preceded the productive vocabulary spurt. Last but not least, the network made a series of under- and over-extension errors *en route* to masterful performance (such as using the word 'dog' exclusively for the family pet or calling all four-legged animals 'dog')—a phenomenon observed in young children using new words (Barrett 1995).

There are several important issues that this model highlights: First, the pattern of behaviour exhibited by the model is highly non-linear *despite the fact that the network architecture and the training environment remain constant throughout learning*. The only changes that occur in the network are small increments in the connections that strengthen the association between an image and its corresponding label. No new mechanisms are needed to explain the vocabulary spurt. Gradual changes within a single learning device are, in principle, capable of explaining this profile of development. McClelland (1989) has made a similar point in the domain of children's developing understanding of weight/distance relations for solving balance beam problems (Siegler 1981).

Second, the model predicts that comprehension precedes production. This in itself is not a particularly radical prediction to make. However, it is an emergent property of the network that was not "designed in" before the model was built. More important is the network's prediction that there should be a non-linearity in the receptive direction, i.e., a vocabulary spurt in comprehension. When the model was first built, there was no indication in the literature as to the precision of this prediction. The prediction has since been shown to be correct (Reznick & Goldfield 1992). This model provides a good example of how a computational model can

be used not only to evaluate hypotheses about the nature of the mechanisms underlying some behaviour but also to generate predictions about the behaviour itself. The ability to generate novel predictions about behaviour is important in simulation work as it offers a way to evaluate the generality of the model in understanding human performance.

The behavioural characteristics of the model are a direct outcome of the interaction of the linguistic and visual representations that are used as inputs to the network. The non-linear profile of development is a direct consequence of the learning process that sets up the link between the linguistic and visual inputs and the asymmetries in production and comprehension can be traced back to the types of representation used for the two types of input. The essence of the interactive nature of the learning process is underscored by the finding that the network learns less quickly when only required to perform the production task. Learning to comprehend object labels at the same time as learning to label objects enables the model to learn the labels faster.

It is important to keep in mind that this simulation is a considerable simplification of the task that the child has to master in acquiring a lexicon. Words are not always presented with their referents and even when they are it is not always obvious (for a child who doesn't know the meaning of the word) what the word refers to. Nevertheless, within the

constraints imposed upon the model, its message is clear: New behaviours don't necessarily require new mechanisms and systems integrating information across modalities can reveal surprising emergent properties that would not have been predicted on the basis of exposure to one modality alone.

Small is Beautiful

The immature state of the developing infant places her at a decided disadvantage in relation to her mature, skilled caregivers. In contrast, the new born of many other species are endowed with precocious skills at birth. Why is homo sapiens not born with a set of cognitive abilities that match the adult of the species? This state of affairs may seem all the more strange given that we grow very few new neurons after birth and even synaptic growth has slowed dramatically by the first birthday. In fact, there may be important computational reasons for favouring a relatively immature brain over a cognitively precocious endowment.

A complete specification of a complex nervous systems would be expensive in genetic resources. The programming required to fully determine the precise connectivity of any adult human brain far exceeds the information capacity in the human genome. Much current research in brain development and developmental neurobiology points to a dramatic genetic underspecification of the detailed architecture of the neural

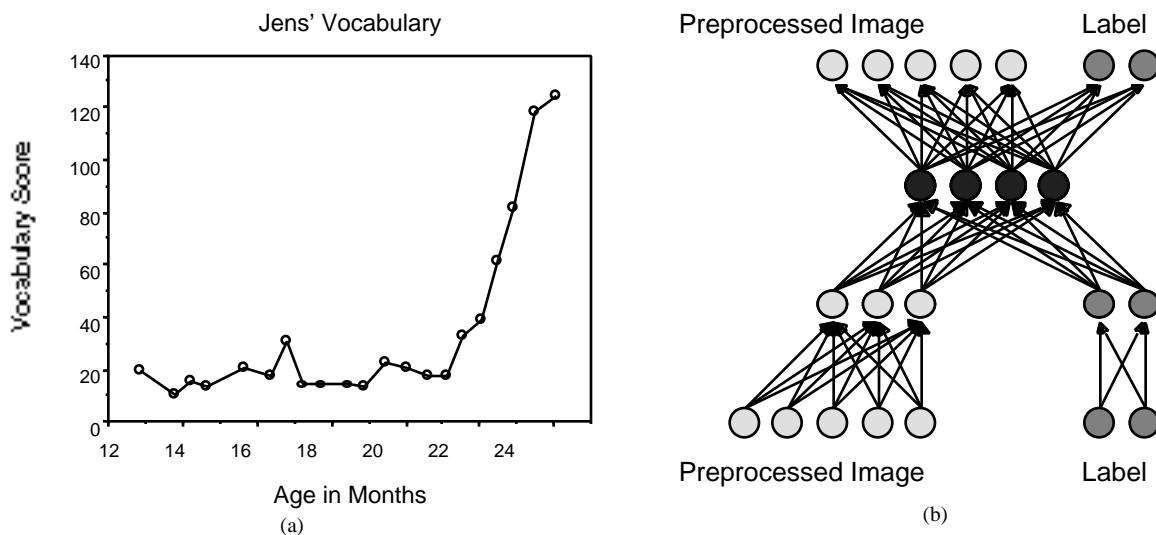


Figure 4: (a) Profile of vocabulary scores typical for many children during their second year—taken from Plunkett (1993). Each data point indicates the number of different words used by the child during a recording session. It is usually assumed that the “bumps” in the curve are due to sampling error, though temporary regressions in vocabulary growth cannot be ruled out. The vocabulary spurt that occurs around 22 months is observed in many children. It usually consists of an increased rate of acquisition of nominals—specifically names for objects (McShane 1979). (b) Simplified version of the network architecture used in Plunkett, Sinha, Møller & Strandsby 1992. The image is filtered through a retinal pre-processor prior to presentation to the network. Labels and images are fed into the network through distinct “sensory” channels. The network is trained to reproduce the input patterns at the output—a process known as auto-association. Production corresponds to producing a label at the output when only an image is presented at the input. Comprehension corresponds to producing an image at the output when only a label is presented at the input.

pathways that characterise the mature human brain—particularly in the neo-cortex. So how does the brain know how to develop? It appears that evolution has hit upon a solution that involves a trade-off between nature and nurture: You don't need to encode in the genes what you can extract from the environment. In other words, use the environment as a depository of information that can be relied upon to drive neural development.

The emergence of neural structures in the brain is entirely dependent upon a complex interaction of the organism's environment and the genes' capacity to express themselves in that environment. This evolutionary engineering trick allows the emergence of a complex neural system with a limited investment in genetic pre-wiring. Of course, this can have disastrous consequences when the environment fails to present itself. On the other hand, the flexibility introduced by genetic underspecification can also be advantageous when things go wrong, such as brain damage. Since information is available in the environment to guide neural development, other brain regions can take over the task of the damaged areas. Underspecification and sensitivity to environmental conditions permit a higher degree of individual specialisation and adaptation to changing living conditions. Starting off with a limited amount of built-in knowledge can therefore be an advantage if you're prepared to take the chance that you can find the missing parts elsewhere.

There are, however, other reasons for wanting to start out life with some limits on processing capacity. It turns out that some complex problems are easier to solve if you first tackle them from a over-simplistic point of view. A good example of this is Elman's (1993) simulation of grammar learning in a simple recurrent network. The network's task was to predict the next word in a sequence of words representing a large number of English-like sentences. These sentences included long distance dependencies, i.e., the sentences included embedded clauses which separated the main noun from the main verb. Since English verbs agree with their subject nouns in number, the network must remember the number of the noun all the way through the embedded clause until it reaches the main verb of the sentence. For example, in a sentence like "The boy with the football that his parents gave him on his birthday chases the dog", the network must remember that "boy" and "chases" agree with each other. This is the type of phenomenon which Chomsky (1959) used to argue against a behaviourist approach to language.

Even after a considerable amount of training, the network did rather poorly at predicting the next word in the sequence—as do humans (cf. "The boy chased the ???"). However, it did rather well at predicting the grammatical category of the next word. For example, it seemed to know when to expect a verb and when to expect a noun, suggesting that it had learnt some fundamental facts about the grammar of the language to which it had been exposed. On the other hand, it did very

badly on long distance agreement phenomena, i.e., it could not predict correctly which form of the verb should be used after an intervening embedded clause. This is a serious flaw if the simulation is taken as a model of grammar learning in English speakers, since English speakers clearly are able to master long-distance agreement.

Elman discovered two solutions to this problem: The network could learn to master long-distance dependencies if the sentences to which it was initially exposed did not contain any embedded clauses and consisted only of sequences in which the main verb and its subject were close together. Once the network had learnt the principle governing subject-verb agreement under these simplified circumstances, embedded clauses could be included in the sentences in the training environment and the network would eventually master the long-distance dependencies. Exposure to a limited sample of the language helped the network to decipher the fundamental principles of the grammar which it could then apply to the more complex problem. This demonstration shows how "motherese" might play a facilitatory role in language learning (Snow 1977).

Elman's second solution was to restrict the memory of the network at the outset of training while keeping the long distance dependencies in the training sentences. The memory constraint made it physically impossible for the network to make predictions about words more than three or four items downstream. This was achieved by resetting the context units in the recurrent network and is equivalent to restricting the system's working memory. When the network was constrained in this fashion it was only able to learn the dependencies between words that occurred close together in a sentence. However, this limitation had the advantage of preventing the network from being distracted by the difficult long-distance dependencies. So again the network was able to learn some of the fundamental principles of the grammar. The working memory of the network was then gradually expanded so that it had an opportunity to learn the long-distance dependencies. Under these conditions, the network succeeded in predicting the correct form of verbs after embedded clauses.

The initial restriction on the system's working memory turned out to have beneficial effects: Somewhat surprisingly, the network succeeded in learning the grammar underlying word sequences when working memory started off small and was gradually expanded, while it failed when a full working memory was made available to the network at the start of training.

The complementary nature of the solutions that Elman discovered to the problem of learning long-distance agreement between verbs and their subjects highlights the way that nature and nurture can be traded off against one another in the search for solutions to complex problems. In

one case, exogenous environmental factors assisted the network in solving the problem. In the other case, endogenous processing factors pointed the way to an answer. In both cases, though, the solution involved an initial simplification in the service of long term gain. In development, big does not necessarily mean better.

Current Shortcomings

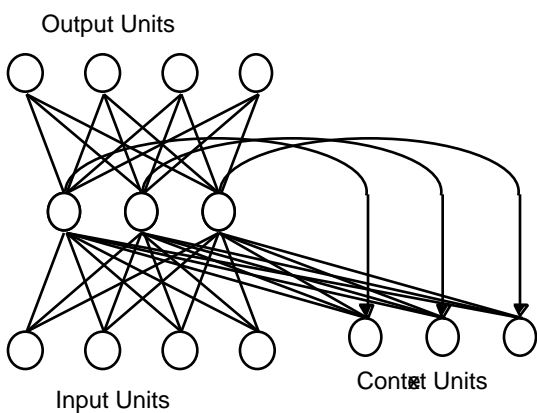
One trial learning

Children and adults learn quickly. For example, a single reference to a novel object as a *wug* may be sufficient for a child to use and understand the term appropriately on all subsequent occasions. The connectionist models described in this paper use learning algorithms which adjust network connections in a gradualistic, continuous fashion. An outcome of this computational strategy is that new learning is slow. To the extent that one trial learning is an important characteristic of human development, these connectionist models fail to provide a sufficiently broad basis for characterising the mechanisms involved in development.

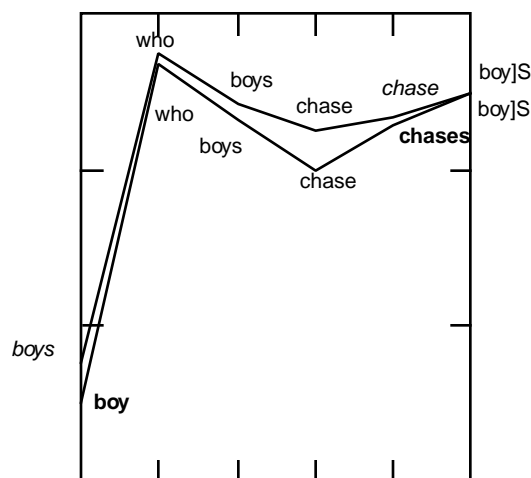
There are two types of solution that connectionist modellers might adopt in response to these problems. First, it should be noted that connectionist learning algorithms are not inherently incapable of one trial learning. The rate of change in the strength of the connections in a network is determined by a parameter called the learning rate. Turning up the learning rate will result in faster learning for a given input pattern. For example, it is quite easy to demonstrate one trial learning in a

network that exploits a Hebbian learning algorithm. However, a side effect of using high learning rates is that individual training patterns can interfere with each other, sometimes resulting in undesirable instabilities in the network. Of course, interference is not always undesirable and may help us explain instabilities in children's performance such as in their acquisition of the English past tense. Generally, though, *catastrophic interference* between training patterns (when training on one pattern completely wipes out the traces of a previously trained pattern) is undesirable. One way to achieve one trial learning without catastrophic interference is to ensure that the training patterns are orthogonal (or dissimilar) to each other. Many models deliberately choose input representations which fulfil this constraint.

An alternative response to the problem of one trial learning in networks is to suggest that in some cases it is illusory, i.e., when individuals demonstrate what is apparently entirely new learning they are really exploiting old knowledge in novel ways. Vygotsky (1962) coined the term the *Zone of Proximal Development* to describe areas of learning where change could occur at a fast pace. Piaget (1952) used the notion of *moderate novelty* in a similar fashion. The performance of networks can change dramatically over just a couple of learning trials. For example, the Plunkett et al. (1992) simulation of vocabulary development exhibited rapid vocabulary growth after a prolonged period of slow lexical learning. The McClelland (1989) balance beam simulation shows similar stage-like performance. In both cases, the networks gradually move towards a state of readiness that then suddenly catapults them into higher levels of behaviour.



(a)



(b)

Figure 5: (a) A simple recurrent network (Elman 1993) is good at making predictions. A sequence of items is presented to the network, one at a time. The network makes a prediction about the identity of the next item in the sequence at the output. Context units provide the network with an internal memory that keeps track of its position in the sequence. If it makes a mistake, the connections in the network are adapted slightly to reduce the error. (b) When the input consists of a sequence of words that make up sentences, the network is able to represent the sequences as trajectories through a state space. Small differences in the trajectories enable the network to keep track of long-distance dependencies

Some one trial learning may be amenable to this kind of analysis. It seems unlikely, however, that all one trial learning is of this kind.

Defining the task and the teacher

Some network models are trained to carry out a specific task that involve a teacher. For example, the Rumelhart & McClelland model of past tense acquisition is taught to produce the past tense form of the verb when exposed to the corresponding stem. These are called supervised learning systems. In these simulations, the modeller must justify the source of the teacher signal and provide a rationale for the task the network is required to perform. Other models use an unsupervised form of learning such as auto-association (Plunkett et al., 1992) or prediction (Elman 1993, Mareschal et al., 1995). In these models, the teacher signal is the input to the network itself. In general, connectionist modellers prefer to use unsupervised learning algorithms. They involve fewer assumptions about the origins of the signal that drive learning. However, some tasks seem to be inherently supervised. For example, learning that a dog is called a *dog* rather than a *chien* involves exposure to appropriate supervision. Nevertheless, it is unclear how the brain goes about conceptualising the nature of the task to be performed and identifying the appropriate supervisory signal. Clearly, different parts of the brain end up doing different types of things. One of the challenges facing developmental connectionists is to understand how neural systems are able to define tasks for themselves in a self-supervisory fashion and to orchestrate the functioning of multiple networks in executing complex behaviour.

Biological plausibility

Throughout this paper I have tried to demonstrate how connectionist models can contribute to our understanding of the mechanisms underlying linguistic and cognitive development. Yet the learning algorithms employed in some of the models described here are assumed to be biologically implausible. For example, backpropagation (Rumelhart, Hinton & Williams 1986) involves propagating error backwards through the layers of nodes in the network. However, there is no evidence indicating that the brain propagates error across layers of neurons in this fashion and some have argued that we are unlikely to find such evidence (Crick 1989).

There is a considerable literature concerning the appropriate level of interpretation of neural network simulations (see Smolensky 1988). For example, it is often argued that connectionist models can be given an entirely functionalist interpretation and the question of their relation to biological neural networks left open for further research. In other words, the vocabulary of connectionist models can be couched at the level of software rather than hardware, much like the classical symbolic approach to cognition. Many developmental connectionists, however, are concerned to understand the

nature of the relationship between cognitive development and changes in brain organisation. Connectionist models which admit the use of biologically implausible components appear to undermine this attempt to understand the biological basis of the mechanisms of change.

Given the success of connectionist approaches to modelling development, it would seem wasteful to throw these simulations onto the waste bin of the biologically implausible. Clearly, the most direct way forward is to implement these models using biologically plausible learning algorithms, such as Hebbian learning. Nevertheless, there are several reasons for tentatively accepting the understanding achieved already through existing models. First, algorithms like backpropagation may not be that implausible. The neurotransmitters that communicate signals across the synaptic gap are still only poorly understood but it is known that they communicate information in both directions. Furthermore, information may be fed backwards through the layered system of neurons in the cortex—perhaps exploiting the little understood back projecting neurons in the process.

A second, related proposal assumes that algorithms like backpropagation belong to a family of learning algorithms, all of which have similar computational properties and some of which have biologically plausible implementations. The study of networks trained with backpropagation could turn out to yield essentially the same results as networks trained with a biologically plausible counterpart. There is some support for this point of view. For example, Plaut & Shallice (1993) lesioned a connectionist network trained with backpropagation and compared its behaviour with a lesioned network originally trained using a contrastive Hebbian learning algorithm. The pattern of results obtained were essentially the same for both networks. This result does not obviate the need to build connectionist models that honour the rapidly expanding body of knowledge relating to brain structure and systems. However, it does suggest that given the rather large pockets of ignorance concerning brain structure and function, we should be careful about jettisoning our hard won understanding of computational systems that may yet prove to be closely related to the biological mechanisms underlying development.

Some Lessons

A commonly held view has been that connectionism involves a *tabula rasa* approach to human learning and development. It is unlikely that any developmental connectionist has ever taken this position. Indeed, it is difficult to imagine what a *tabula rasa* connectionist network might look like. All the models reviewed in this article assume a good deal of built-in architectural and processing constraints to get learning off the ground. In some cases, such as the Rumelhart & McClelland model of the past tense, the initial constraints are quite

modest. In others, such as the Mareschal et al., model of visual tracking and reaching, the initial architectural and computational assumptions are rather complex. These modelling assumptions, together with the task definition, imply a commitment to the ingredients that are necessary to get learning off the ground.

What is needed to get learning off the ground? We have seen that there are two main sources of constraint:

1. The initial state of the organism embodies a variety of architectural and computational constraints that determine its information processing capabilities.
2. Environmental structure supports the construction of new representational capacities not initially present in the organism itself.

Modelling enables us to determine whether a theory about the initial state of the organism can make the journey to the mature state given a well-defined training environment. Modelling also enables us to investigate the minimal assumptions about the initial state that are needed to make this journey.

A minimalist strategy may not necessarily provide an accurate picture of the actual brain mechanisms that underlie human development. However, it provides an important potential contrast to theories of the initial state that are based on arguments from the poverty of the stimulus. Investigating the richness of the stimulus shifts the burden away from the need to postulate highly complex, hard-wired information processing structures. A minimalist strategy may also provide valuable insights into alternative solutions that the brain may adopt when richer resources fail.

Theories about the initial state of the organism cannot be dissociated from theories about what constitutes the organism's effective environment. Release two otherwise identical organisms in radically different environments and the representations they learn can be quite disparate. Connectionist modelling offers an invaluable tool for investigating these differences as well as examining the necessary conditions that permit the development of the emergent representations that we all share.

Acknowledgements

This manuscript was produced while the author was engaged in a collaborative book project together with Jeff Elman, Liz Bates, Mark Johnson, Annette Karmiloff-Smith and Domenico Parisi. The content of this manuscript has been influenced profoundly by discussions with my book co-authors. The reader is strongly recommended to consult Elman et al. (In press) for a more wide-ranging and detailed discussion of the issues raised here.

References

- Baillargeon, R. (1993). The object concept revisited: New directions in the investigation of infant's physical knowledge. In: C. E. Granrud (Ed.), *Visual perception and cognition in infancy*, 265-315. London, UK: LEA.
- Barrett, M. D. (1995). Early Lexical Development. In P. Fletcher & B. MacWhinney (Eds.), *The Handbook of Child Language*, (pp. 362-392). Oxford: Blackwells.
- Bates, E., Bretherton, I., & Snyder, L. (1988). *From First Words to Grammar: Individual Differences and Dissociable Mechanisms*. Cambridge, MA: Cambridge University Press.
- Bliss, T. V. P., & Lømo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology*, **232**, 331-356.
- Chomsky, N. (1959). Review of Skinner's verbal behavior. *Language*, **35**, 26-58.
- Cottrell, G. W., & Plunkett, K. (1994). Acquiring the mapping from meanings to sounds. *Connection Science*, **6**(4), 379-412.
- Crick, F. H. C. (1989). The real excitement about neural networks. *Nature*, **337**, 129-132.
- Elman, J. L. (1993). Learning and development in neural networks: the importance of starting small. *Cognition*, **48**(1), 71-99.
- Elman, J., Bates, E., Karmiloff-Smith, A., Johnson, M., Parisi, D., & Plunkett, K. (In press). *Rethinking Innateness: Development in a connectionist perspective*. Cambridge, MA: MIT Press.
- Foldiak, P. (1991). Learning invariance in transformational sequences. *Neural Computation*, **3**, 194-200
- Karmiloff-Smith, A. (1979). Micro- and macrodevelopmental changes in language acquisition and other representational systems. *Cognitive Science*, **3**, 91-118.
- MacWhinney, B. & Leinbach, A. J. (1991) Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, **40**, 121-157.
- McClelland, J. L. (1989). Parallel distributed processing: implications for cognition and development. In R. G. M. Morris (Ed.), *Parallel Distributed Processing: Implications for Psychology and Neurobiology*. Oxford: Clarendon Press.
- McShane, J. (1979). The development of naming. *Linguistics*, **17**, 879-905.
- Marchman, V. A. (1993). Constraints on Plasticity in a Connectionist Model of the English Past Tense. *Journal of Cognitive Neuroscience*, **5**(2), 215-24.
- Marcus, G. F., Ullman, M., Pinker, S., Hollander, M., Rosen, T. J. & Xu, F. (1992) Overregularization in language acquisition. Monographs of the Society for Research in Child Development, **57**(4), Serial No. 228.
- Mareschal, D., Plunkett, K., & Harris, P. (1995). Developing Object Permanence: A Connectionist Model. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*, (pp. 170-175). Mahwah, NJ.: Lawrence Erlbaum Associates.

- Piaget, J. (1952). *The Origins of Intelligence in the Child*. New York: International Universities Press.
- Piaget, J. (1955). Les stades du développement intellectuel de l'enfant et de l'adolescent. In P. O. e. al. (Ed.), *Le probleme des stades en psychologie de l'enfant*. Paris: Presses Univer. France.
- Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a Parallel Distributed Processing Model of language acquisition. *Cognition*, **29**, 73-193.
- Plaut, D. C., & Shallice, T. (1993). Deep Dyslexia: A Case Study of Connectionist Neuropsychology. *Cognitive Neuropsychology*, **10**(5), 377-500.
- Plunkett, K. (1995). Connectionist Approaches to Language Acquisition. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of Child Language*, (pp. 36-72). Oxford: Blackwells.
- Plunkett, K. & Marchman, V. (1991) U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition*, **38**, 43-102.
- Plunkett, K. & Marchman, V. (1993) From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, **48**, 1-49.
- Plunkett, K., Sinha, C. G., Møller, M. F. & Strandsby (1992) Symbol grounding or the emergence of symbols? Vocabulary growth in children and a connectionist net. *Connection Science*, **4**, 293-312.
- Reznick, J. S. & Goldfield, B. A. (1992) Rapid change in lexical development in comprehension and production. *Developmental Psychology*, **28**, 406-413.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the Microstructure of Cognition*, Vol 1: Foundations, (pp. 318-362.). Cambridge, MA: MIT Press.
- Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tense of English verbs. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: explorations in the microstructure of cognition*. Cambridge: MIT Press.
- Siegler, R. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development*, **46**, Whole No. 2.
- Snow, C. E. (1977). Mothers' speech research: From input to interaction. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition*. Cambridge: Cambridge University Press.
- Spelke, E. S., Katz, G., Purcell, S. E., Ehrlich, S. M. & Breinlinger, K. (1994) Early knowledge of object motion: continuity and inertia. *Cognition*, **51**, 131-176.
- von Hofsten, C (1989). Transition mechanisms in sensorimotor development. In: A. de Ribaupierre (Ed.), *Transition mechanisms in child development: The longitudinal perspective*, 223-259. Cambridge, UK: Cambridge University Press.
- Vygotsky, L. (1962). *Thought and language*. Cambridge: MIT Press.
- Ungerlieder, L. G. Mishkin, M. (1982). Two cortical visual systems. In: D. J. Ingle, M. A. Goodale, & Mansfield (Eds.), *Analysis of visual behavior*. Cambridge, MA: MIT Press.
- van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, **98**, 3-53.