In Child Language by M. Alderidge (Eds.). Multilingual Matters Ltd, Clevedon, 1997

Simulation of 'Sensori-motor Stage V and VI' Language Development: A Connectionist Network Approach

Syed Sibte Raza Abidi and Khurshid Ahmad
University of Surrey, Guildford.

Abstract

C*onnectionist networks*, or the so-called *neural networks*, provide a basis for studying child language development in that these networks emphasise learning, either from observations or from being *told,* and that the design of these networks simplifies questions related to the representation of linguistic and world knowledge through the use of a network of 'simple' nodes and links. We report a connectionist simulation of this phenomenon focusing on the transition from one-word to two-word language. The connectionist simulation comprises the simulation of *concept memory*, *word lexicon*, *semantic and conceptual relations* and *word-order*. The data used in the simulation has its origins in the longitudinal psycholinguistic study of infants through their various stages of cognitive development, including sensori-motor (stages V and VI) and pre-operational stages. Bloom's (1973) archives of child language data was used in 'training' the connectionist networks. The concept representation scheme for simulating a child's concept memory is semantic feature oriented (Katherine Nelson, 1973) and the semantic relations between concepts are based on Roger Brown's (1973) analysis. The connectionist simulation was carried out using ACCLAIM - *A Connectionist Child LAnguage Development and Imitation Model*. ACCLAIM is a hybrid connectionist architecture comprising 'supervised' and 'unsupervised' learning connectionist networks, and takes into account the diverse nature of inputs to and outputs from a child learning language.

1. Introduction

Learning is a much debated topic in artificial intelligence (AI), neurobiology and linguistics. Assuming that language is unique to human beings, the so-called natural language, then it can be argued that the development of language amongst children can provide us with pointers to a number of open questions in the literature. For instance, the role of the environment in determining motor and cognitive development has been studied extensively in developmental psychology and neurobiology. The studies in child language development are at once an exemplar of and an inspiration to workers in human learning. For us, child language is an exemplar of human learning, a development that has been studied from time immemorial, has a body of systematically collected data, and is an aspect of human development that has thus far eluded a well-grounded, objective theoretical framework. We believe that child language development can benefit from the objectivity that may be implicit in connectionist network methodology.

Child language development theories have motivated the collection of substantial amounts of observational data. The data collection exercises involve a number of interesting hypotheses about how a child may perceive the 'world' around him/herself. This may include his/her beliefs, desires and feelings. In order to express such internal states, Bloom has argued that language development needs to account for the development of concepts, of lexica, of semantics, of syntax and the development of discourse (1993: 96). Therefore, it is possible to argue that if one were successful in synthesising psycholinguistic observations, particularly during the various stages of language development: spanning the onset of language, (c. 9-12 months), the vocabulary spurt, and the transition to multi-word speech (c. 18-24 months), with the developmental neurobiological observations, then one can have a psychologically plausible and neurobiologically tangible description of the language of children.

Advances in neurobiology have inspired computing scientists to build systems that crudely mimic the organisation of the brain: the so-called *neural networks* or *connectionist networks* are potentially parallel processing systems, involving co-operative computations among locally connected processing units. Their crudity notwithstanding, and the fact that much of what inspired the early neural networks pioneers were simplistic notions in behaviourism, neural networks have one important characteristic that makes them more plausible than, say, artificially intelligent systems or conventional procedural computing systems. This characteristic is their ability to 'learn' from observations and from stimuli over a period of time; learning is inherent in the design of connectionist networks, whereas both the conventional and AI systems are incapable of such learning. Bechtal and Abrahamsen have argued that 'connectionism could be viewed as a modern mechanism for achieving stage-like states by means of the heretofore somewhat mysterious processes of accommodation and assimilation' (1991: 271).

The simulation of physical and biological systems, based generally on mathematical models, has been of considerable pedagogic and theoretical value in physics, chemistry, and perhaps to a lesser extent, in biology. The simulations of problem-solving behaviour in artificial intelligence, though not overtly mathematical yet still algorithmic, have led to the explication of problem-solving behaviour. The same could perhaps be said about the developments in computational linguistics, particularly the development of programs that syntactically analyse phrases and sentences. The success of simulation-based studies have inspired a number of workers including Siskind (1990), MacWhinney (1987), Hill (1983), Langley (1982) and Selfridge (1982), to build simulation models of child language development. These studies were either based on procedural methods and techniques used in conventional computer science or on methods and techniques in artificial intelligence.

Such simulations, though very instructive, cannot be used to explicate much about learning in that the conventional computing and AI systems used in child language simulation studies do not have any learning mechanisms. And, in any case, the focus of these studies was on procedural aspects of human problem-solving. Connectionist simulation appears more promising in the context of child language development; through its underpinnings in neurobiology and psychology, and the learning mechanisms that are inherent in the design of connectionist networks will perhaps allow a better simulation of child language development.

We have developed ACCLAIM - A Connectionist Child LAnguage Development & Imitation Model to simulate child language development within the age group 9-24 months. ACCLAIM is a 'hybrid' connectionist architecture implementing a variety of connectionist networks, including *Kohonen maps*, *backpropagation networks*, *additive Grossberg Networks*, networks with *Hebbian connections* incorporating the *spreading activation* mechanism. ACCLAIM has been used to simulate the development of concepts amongst children together with the lexicalisation of these concepts: the *concept memory* and *word lexicon* have been simulated using Kohonen maps and are linked together through a Hebbian connection based *concept lexicalisation* network. Backpropagation networks have been used to implement a *conceptual relation* network (for one-word sentences) and a *word-order* network (for two-word sentences). Children's evolving 'semantic' performance has been simulated using additive Grossberg network. Thus, aspects of what can be construed to be innate development have been simulated using unsupervised learning regimes, like Kohonen maps and Hebbian connections, and environmentally-determined features of language development have been simulated using supervised learning regimes, like backpropagation networks. ACCLAIM has been trained on 'realistic' child language data and has learnt to recognise and produce one-word and two-word sentences.

Our work is different from the earlier simulations of child language development in three significant respects: first, we simulate a number of aspects of human language including lexical organisation and lexical access, conceptual memory, semantics, pivot grammar and word order for studying evolving linguistic behaviour. Second, our focus is on the development or the evolution of linguistic behaviour amongst children. The notions of innate structures notwithstanding, language is learnt over a period of time and involves environmental input. This includes input from the physical environment, caretakers, siblings and others, together with language learnt by the child on his or her own initiative with or without supervision, either through the maturity of the nervous system or through some other natural gift. Third, we believe that the interdependence of language learnt via environmental input and self-motivated language learning, can surely influence the kinds of connectionist network architectures that either simulate 'supervised learning' or 'unsupervised learning'.

In this paper we limit the discussion to the simulation of the development of the connectionist 'concept memory' which itself indicates how connectionism provides opportunities for operationalising child language theories on concept representation and development. We start with a brief introduction of connectionism and, the inherent learning mechanism, and a description of the connectionist network - Kohonen map used to implement the concept memory (Section 2). In Section 3 we present details of the connectionist architecture of ACCLAIM. In Section 4 we discuss the nature of children's concepts and present a viable representation scheme based on notions suggested by Katherine Nelson (1973), Lois Bloom (1973) and others. Our connectionist simulation of concept development is based on the proposed concept representation scheme. Section 5 describes a connectionist simulation of the development of the concept memory. In Section 6 we give a brief description of the simulation of a two-word child-like sentence production.

2. Connectionism: A brief introduction

Connectionism is a research discipline that aims to understand the nature of human intelligence by simulating aspects of human behaviour through a collection of idealised neurons. Connectionism draws much of its inspiration from neurosciences in that the *neuron* is taken as the basic *processing unit*. Each such processing unit is characterised by an *activation level* (analogous to the state of polarisation of a neuron), an *output value* (representing the firing rate of the neuron), and a set of *input* and *output connections* (representing the neuron's axons and dendrites) from and to other units, respectively. These characteristics are expressed in a mathematical formalism such that a unit's activation level and output value are expressed as (real) numbers, and its connections with other units have an associated *weight* (synaptic strength) which determine the effect of the incoming input on the activation level of the unit. The processing units are provided with a variety of 'stimuli' and are expected to 'respond' in a manner that mimics aspects of human behaviour.

The use of a collection of idealised neurons is at the heart of studies in 'microcognition'- a term coined in the seventies. This term emphasises the role of simple individual processors or processing units, rather than the generic approach of symbol manipulation as is generally practised in AI.

Connectionist networks mimic the neural structure of the brain rather simplistically in that a connectionist network comprises a large number of computationally simple processing units which are highly interconnected through *plastic* connections. The 'plasticity' in the architecture of a connectionist network is introduced with the help of varying connection weights that can change over time and with experience. The configuration of the processing units dynamically adapts to the environment as a consequence of 'learning'. Put simply, learning in connectionist networks can be envisaged as the problem of finding a set of connection weights which allow the connectionist network to store experiential knowledge and to exploit it to simulate the desired behaviour. One can then argue that connectionist networks have a 'natural' propensity for storing experiential knowledge which is acquired and retained through 'training' or 'learning' as opposed to explicit programming. Computationally, connectionism emphasises parallelism, distributed control and the plasticity of connections between processing units that comprise the parallel architecture. The key notions in connectionism include learning through experience and learning through evolution.

2.1. 'Learning' in a connectionist network

Learning in connectionist networks has substantial resonances with the work of behavioural psychologists during the 1940's and 1950's. Textbooks on connectionist networks begin with statements like 'learning would involve relatively enduring changes in a system of given architecture that results from its interaction with the environment. The most obvious form of learning is adjustment in the weights of connections' (Bechtal and Abrahamsen, 1991: 270).

Learning is effected through changes in the strength of connections between individual processing units in a connectionist network. Put simply, given a set of inputs x_1 .... x_n (a vector symbolically denoted as X) to a system, the system generates a set of outputs, y_1 ... y_n denoted symbolically as (a vector) Y. This is achieved computationally by relating X and Y through a matrix of connection weights w_{11} ... w_{nn}, denoted as W. This interrelationship matrix assumes that one, some or all the inputs influence individual outputs (see Figure 1):

$y_1 = w_{11}x_1 + w_{12}x_2 + ... + w_{1n}x_n$; $y_n = w_{n1}x_1 + w_{n2}x_2 + ... + w_{nn}x_n$ (and similarly for y_2 .... y_{n-1}).

Learning in the above simplification is then the change of the weights W.

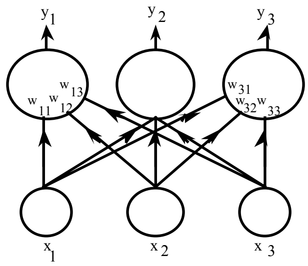

Figure 1: A connectionist network, showing processing units and connections

Connectionist networks can be trained in a number of ways, but generally three different types of learning mechanisms are more popular in the literature (Caudill and Butler, 1992):

> **Supervised Learning: The network is provided with an input pattern along with a desired output pattern. The learning law for such networks typically computes an *error*, that is the difference between the desired output of the network to its actual output. The computed error is then used to modify the interconnections between the units. In supervised learning connectionist networks, 'certain output nodes are trained to respond to certain "exemplar" patterns, and the changes in connection weights due to learning cause those same nodes to respond to more general classes of patterns' (Levine, 1991: 196). Best exemplars of supervised learning are perceptrons and backpropagation networks (Rumelhart et al, 1986).**
>
> **Graded Learning: Similar to supervised training except that the exact desired output is not provided, only a 'grade' on how well the network is learning is specified.**
>
> **Unsupervised learning: The network is presented only with a series of input patterns and is given no information or feedback at all about its performance or desired output. Such training procedures are generally used for categorisation or statistical modelling applications because the network's response cannot be predicted by the designer of the network. Within unsupervised connectionist networks 'input patterns are presented in some sequence and the network discovers through self-organisation[1] a "natural" categorisation of the sensory world' (Levine, 1991: 196). Best exemplars of unsupervised learning connectionist networks are Kohonen maps, competitive networks and Hebbian learning networks.**

2.2. Connectionism and Piagetian notions of learning

Some connectionists and philosophers of science have reinterpreted Piagetian notions of learning in the connectionist paradigm. Specifically, this reinterpretation focuses on Piagetian notions of *accommodation* and *assimilation*. For instance, McClelland's essay on the implications of connectionism for 'cognition and development' includes the description of a 'learning principle' governing cognitive development: 'adjust the parameters of the mind in proportion to the extent to which their adjustment can produce a reduction in the discrepancy between expected and observed events' (1989: 20). McClelland further notes that this learning principle captures the 'residue of Piaget's accommodation process', in that accommodation involves an adjustment of mental structures in response to discrepancies between an 'expected' and an 'observed' event. For McClelland, the novelty of this principle is that it can be implemented using a connectionist network. Bechtal, a philosopher of science, and Abrahamsen, a developmental psychologist, have discussed how connectionism may help in reinterpreting 'certain Piagetian constructs': assimilation is reinterpreted in terms of 'the tendency of an interactive network to settle into the most appropriate of its stable states when input is presented to it; in Piaget's language this is the schema to which the experience has been assimilated'. Similarly, accommodation is reinterpreted as 'the changes in activations as well as weights that occur in order to assimilate the experience' (1991: 271). Indeed, Bechtal and Abrahamsen have claimed that not only can one reinterpret Piagetian constructs, but it might

---

[1]An artificial intelligence program that is able to modify itself by adapting to its environment without help from an outsider. It is able to profit from experience. (Mercadal 1990:257)

be also possible to augment and replace some of those constructs! Table 1 compares Piagetian constructs suggested by McClelland with notions in connectionism.

Piagetian Constructs	Analogous Connectionist Notions
Parameters of the mind	Connections among units. Both entities are amenable to alteration due to experience.
Expected event	Desired pattern of activation over the network's output units.
Observed event	Actual pattern of activation produced over the network's output units.
Adjustment of the parameters	Connectionist learning processes that involve adjustment of connections.
Discrepancy reduction	'Error minimisation' process during connectionist learning, reducing error between expected and observed pattern of activation.

Table 1: Correspondence between Piagetian constructs and analogous connectionist notions.

2.3. Connectionist 'unsupervised learning'

The major simulation discussed in this paper concerns the simulation of how the so-called concept memory develops amongst children. The simulation is based on a connectionist network - Kohonen map which employs an 'unsupervised learning' algorithm. Learning in an unsupervised manner is a recent trend in the connectionist community and is characterised by the fact that it does not rely on the feedback of an external 'teacher' (as is the case with supervised learning) verifying the goodness of learning. Rather, unsupervised connectionist networks learn on their own without any explicit supervision; a type of learning which is seemingly more akin to some aspects of learning observed in a developing child. Unsupervised learning is essentially a self-guided process of 'feature detection' on the part of the connectionist network which involves a discrimination between the different features of the input patterns. Given a set of input patterns (the so-called environment of the connectionist network), learning in unsupervised connectionist networks is accomplished by discovering statistical regularities in the input data (involving the computation of a distance measure - 'Euclidian distance'). Learning also involves establishing relationships between common features in various input patterns which leads to the grouping or the so-called 'categorisation' of similar input patterns. Consider below (Table 2) the unsupervised learning of a few concepts represented as a 5-dimensional feature vector. In connectionist terms, each feature in Table 2 can be implemented as an individual processing unit (an input unit), where the value 1 indicates the presence of the feature and the value 0 indicates its absence.

Concept	Furry Coat?	Has Feathers?	Has Tail?	Four Legs?	Flies?	Is Pet?
Dog	1	0	1	1	0	1
Cat	1	0	1	1	0	1
Horse	0	0	1	1	0	0
Robin	0	1	0	0	1	0
Canary	0	1	0	0	1	1

Table 2: Concepts represented in terms of binary valued features.

To a human it is a trivial task to learn the constituent features of each concept and to group the concepts on the basis of feature similarity. For instance, the concepts 'dog' and 'cat' are categorised as pet animals; 'robin' and 'canary' are both birds; 'horse' is an animal but not a pet. Again this is a much simpler task for a supervised learning connectionist network as the 'teacher' guides the connectionist network to learn the correct grouping of the concepts. However, a connectionist network, say a Kohonen map, employing an unsupervised learning mechanism just relies on its 'innate' ability (statistical mechanisms) to detect common features across the range of input patterns. The Kohonen map learns the concepts in terms of the 'detected' features and regularities in the features are manifested by the grouping or 'categorisation' of similar concepts.

Unsupervised learning is based on a process of 'competition' amongst processing units. Consider a connectionist network which constitutes a five-dimensional (5 units - one unit each for one feature in Table 2) *input layer*. When an input pattern is presented to the input layer of a connectionist network, it projects itself on an *output layer* (consisting of n output units: n = 9), such that the input pattern, i.e. a concept, is represented by a unique output unit. To learn the concepts, the various output units *compete* to represent the input pattern, such that the unit acquiring the highest activation level *wins* the competition and is deemed to represent that particular input pattern. During learning, the connectionist network is configured in such a manner that each concept is represented by a unique output unit. Also, the connectionist network autonomously learns which concepts belong to the same category and thus places similar concepts in proximity in the output layer. Later, we further explicate unsupervised learning in connectionist networks through a brief introduction of the structure and dynamics of a typical unsupervised learning connectionist network - Kohonen map. Trevo Kohonen (1984), a leading connectionist researcher with a keen interest in unsupervised learning mechanisms has proposed a connectionist architecture - Kohonen maps based on the theory of *self-organising feature maps.* This connectionist architecture appears intuitively useful for organising and categorising complex information. The basic tenets of the Kohonen map are as follows:

Structure: A Kohonen map (Figure 2) consists of two distinct layers of processing units: an *input layer* and an *output layer*. The input layer is used to present an input vector to the Kohonen map, and consists of n number of units, where n is the number of features in the (n dimensional) input vector X, where $X = [x_1, x_2, .. \ x_n]$. The output layer, usually referred to as a 'two dimensional map', consists of m number of units arranged in a two-dimensional format, i.e. rows and columns. The output layer maps the n-dimensional input vector to a lower (two) dimensional representation. Both layers are connected by weighted connections, such that each output layer unit is connected to all input layer units. Associated with each output unit, o_i , is an n-dimensional weight vector W that stores the strength of the connections from the output unit o_i to all units in the input layer. The weight vector for unit i would be given as $W_j = [w_{1j}, w_{2j}, ... \ w_{nj}]$.

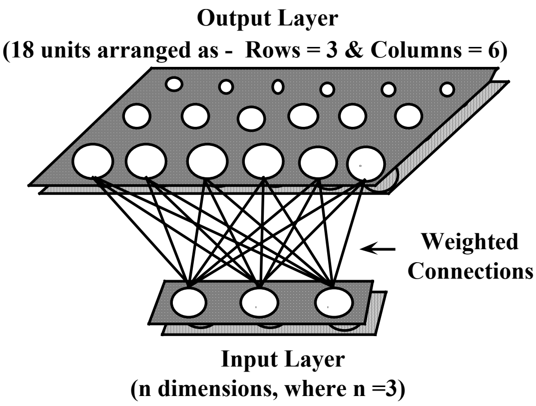

Figure 2: A Kohonen map connectionist network

Network initialisation: Before learning, a Kohonen map is initialised to ensure that it does not contain any prior information. Initialisation is achieved by assigning a random weight value in the range of 0-1 to the components of the weight vectors of all competitive units, resulting in random weighted connections between the input and output layers.

Input presentation: Learning in Kohonen maps is carried out over a number of iterations. In each iteration, an input pattern is randomly chosen from the ensemble of input patterns and presented to the input layer of the Kohonen map. This random selection of the input patterns ensures that the learning taking place does not observe a pre-determined course and is also not biased in any way.

The 'self-organising' learning algorithm: The Kohonen map's learning algorithm is based on a process of 'self-organisation', which changes the connection weights between the input and output layers. The Kohonen map's self-organising learning algorithm comprises the following steps:

i. <u>Initialise</u> the Kohonen map. Initialisation ensures that each competitive unit has a unique weight vector so that no similar topological regions or categories may initially exist.

ii. <u>Present</u> an input pattern to the input layer of the Kohonen map. In each iteration an input pattern is randomly chosen from the set of input patterns.

iii. <u>Determine</u> the competitive unit which best matches the input pattern. This is achieved by firstly calculating for all units the 'Euclidian Distance' between the n-dimensional input patterns and the weight vector of each unit. The unit with the least Euclidian Distance is regarded as the 'image unit' representing the input pattern.

iv. <u>Make</u> the 'image unit' more representative of the input pattern by moving its weight vector closer to the input pattern. As learning progresses, the weight vector of the image unit moves closer and closer to the input pattern.

v. Repeat steps ii - iv for a number of iterations, where in each iteration a different input pattern is randomly chosen from the entire set of input patterns.

The effect of the self-organising learning algorithm is that with increasing iterations, the Kohonen map incrementally learns the input patterns. Learning is to be continued until certain learning criteria are satisfied. At the end of the learning sequence, the input patterns are learnt and represented by the output layer such that similar input patterns are placed in proximity.

The learning criteria: In a Kohonen map each learnt concept, feature, word or any other item of knowledge is represented by a unique unit known as its 'image unit'. In connectionist terms, any information, say a concept, is assumed to be learnt when: (a) the activation level of its image unit is the highest amongst all other units and is approaching unity; and (b) its Euclidean Distance (ED) is minimal, i.e. close to zero. As learning progresses, the image unit of a particular input pattern has its ED minimised by the self-organisation process over a number of iterations, whereas on a reciprocal basis its activation level is increased.

3. A Connectionist Child LAnguage Development & Imitation Model - (ACCLAIM)

In this section we present ACCLAIM - A Connectionist Child LAnguage Development & Imitation Model that was developed to simulate child language development within the age group 9 - 24 months.

Child language development can be characterised by the development of various language related aspects such as concepts, words, semantic relations, word order, incorporation of diverse input stimuli, involvement of different learning mechanisms, broadly categorised as supervised and unsupervised, and an interaction among various 'sub-tasks' including concept development, word learning, concept lexicalisation, concept categorisation, learning of semantic relations and understanding of word order.

Consider the following model of child language development: (i) The child receives two kinds of input from its environment: one is perceptual input that enables the child to categorise entities and events and the other is linguistic input in the form of the caretaker's language, mainly in terms of 'two-word collocates'; (ii) The 'innate' ability of the brain then helps the child to understand his/her environment; 'abstracting' critical semantic features to form concepts and storing them in a 'concept memory'; (iii) Also, the child discriminates between the phonetic content of the linguistic input from caretakers to develop a repertoire of words - 'word lexicon'; (iv) At the end of the sensori-motor development the child learns functional words or 'conceptual relations', and learns to use them as single-word utterances, each in different situations that have common contexts; (v) Furthermore, the child 'learns' to associate concepts and words in an unsupervised manner; (vi) The child generalises further and creates the so-called conceptual categories leading to the development of 'semantic relations' among conceptual categories; (vii) Finally, the child, through a process of trial and error, builds up collocates that conform to the word order in his/her caretakers' language, leading to the production of child-like two-word sentences.

The above-mentioned 'processes' can be simulated by individual connectionist networks or 'connectionist modules'. Some of these processes can be simulated by the so-called connectionist 'supervised learning' algorithms whilst others can be simulated by the use of 'unsupervised learning' algorithms. ACCLAIM was developed by organising independent connectionist networks in a psycholinguistically plausible manner to yield a hybrid of various connectionist networks. Figure 3 shows the 'hybrid' architecture of ACCLAIM.

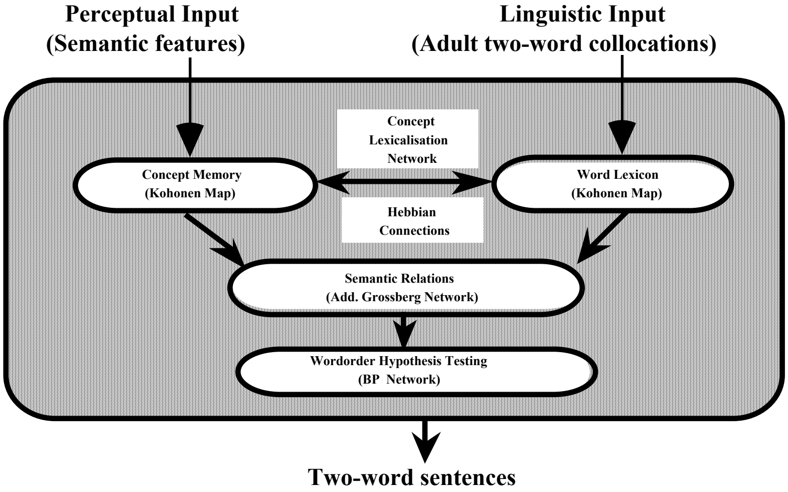

Figure 3: The connectionist architecture of ACCLAIM

In ACCLAIM all the simulations are carried out in a 'developmental' manner: starting with no *a priori* information, the connectionist networks are exposed to a set of input stimuli - 'training patterns'. Over a period of time (iterations) the connectionist networks learn the input stimuli and at the end of the simulation acquire the required 'intellectual status'. 'Environmental' influence during learning is demonstrated by the adaptability of the 'plastic' structure of the connectionist networks to account for information received from the environment. Both the 'developmental' and 'environmental' issues predicate the connectionist implementation of Piaget's learning mechanisms of 'assimilation' and 'accommodation', whereby an initially random connectionist network is transformed into a highly structured connectionist network, storing the information provided to it.

4. Children's concepts: A representation for a connectionist simulation

The child can be considered as 'an active information processor of its environment, encoding features of the world perceived. To encode a representation of the world, the child has to attend to, perceive, and store in memory observations made about the prevalent information. Some of the information attained may be perceptual (immediately apparent information - movement, shape, colour), conceptual (derived information - existence, permanence), speech (adult conversations, television), positive or negative notions (the presence or absence of objects, people or events)' (Nelson, 1973a: 2).

It appears that the child must process information in terms of salient (semantic) features present in his/her environment to form concepts. To represent concepts in a connectionist environment we have adopted a conventional 'semantic feature' based formalism which describes the similarities and differences between various concepts and also helps in defining categories. Each concept in our connectionist representation scheme is represented by a 20-dimensional 'semantic feature vector' comprising two types of features: 'defining features' - determining a category structure, and 'individual features' - distinguishing individual concepts within a category. We discuss below how these defining and individual features are used to construct a semantic feature vector for representing a concept.

The defining features of concepts in our simulation are based on an 'object-oriented' taxonomy - a hierarchical tree distinguishing features at each level. This particular taxonomy was suggested by Nelson (1973). It appears that all children basically distinguish between objects and non-objects. This antonymy is the basis of a 'semantic' tree (Figure 4) : the objects and non-objects are leaves on a tree. Nelson's 'semantic structure' classifies or categorises 'objects' and 'non-objects' at a considerable level of detail, enabling us to determine the category for the object/non-object concept the child may be talking about.

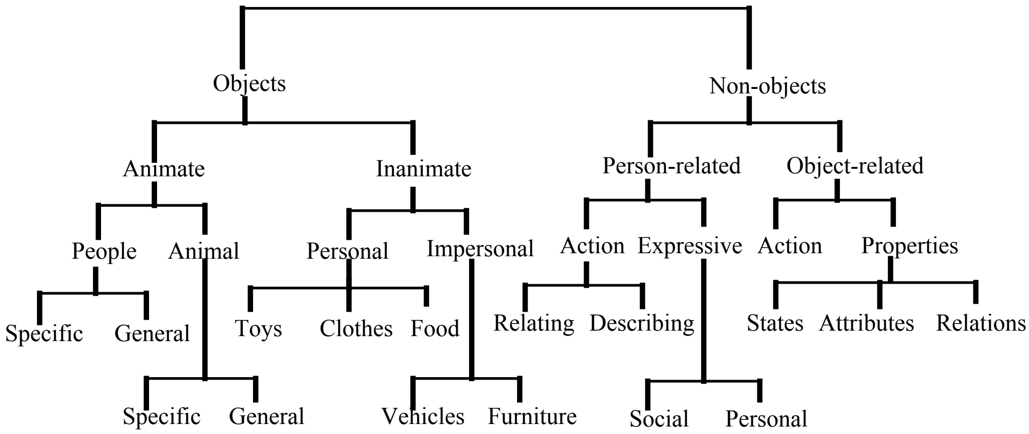

Figure 4: An exemplar semantic structure (Nelson, 1973a: p 119)

For representing various objects and non-objects a child might encounter, we have labelled Nelson's hierarchical structures in terms of binary digits (1 and 0). Two categories at the same level of the tree, for instance 'objects' and 'non-objects', are assigned the values 1 and 0, respectively. Similarly, for the category 'object', the sub-category 'animate' is assigned the value 1 and 'inanimate' is assigned the value 0. Once the tree is labelled, one can determine the 'defining features' for a concept by using the value '1' to indicate that the object/non-object contains a particular feature and '0' to indicate otherwise. Therefore, according to this representation scheme, concepts belonging to the *specific(1)-people(1)-animate(1)-object(1)* category, for instance 'dad', are labelled [1 1 1 1]. Similarly, other conceptual categories are given their own labels, for instance the *generic(0)-people(1)-animate(1)-object(1)* category may be labelled [1 1 1 0]. The vector notation for the two above-mentioned categories is [1,1,1,1] and [1,1,1,0] respectively.

Nelson's discussion is really at a meta-level in the sense that she talks about semantic categories, but there are no individual 'concepts' on this tree. The semantic structure is then useful for determining a set of 'defining features' that categorise objects into various concept categories. Bloom (1973) has argued that children going through the developmental 'stages V and VI' do have access to a number of concepts. Children's possession of a variety of concepts, differing from one another in terms of salient features, suggests that a category level abstraction alone may not suffice to represent children's concepts. Rather, individual concepts need to be analysed in more detail so that it becomes possible to further identify the individual concepts within concept categories. We argue that the features unique to a concept, i.e., the so-called 'individual features' help discriminate one concept from other concepts having the same 'defining features'.

What seems relevant here is the specification of these so-called 'individual features'. Notwithstanding the juxtaposition of philosophical ideas about 'semantic features', we have collected a number of meaningful 'individual features' from various studies reported in child language literature. For instance, children are believed to distinguish various objects by observing aspects such as 'size', 'shape', 'colour' and even, at times, their 'function'. We have examined whether such aspects can be treated as individual (semantic) features that in turn can describe a concept. The 'individual features' which we have collected adequately distinguish various concepts and hence serve the purpose of our connectionist simulation of the development of the concept memory.

In our representation scheme the individual features are based on a taxonomy of children's concepts suggested by Bloom (1973) and recently commented on by Anisfeld (1989). The taxonomy consists of seven different categories: *objects*, *agents*, *events*, *states*, *locations*, *prepositions* and *'function words'* [2]. Each category comprises a number of 'individual features' that we believe may represent the concepts associated with the category. In Figure 5, we show a cross-section of the 'individual feature tree' for the category 'agents', mainly focusing on features related to 'Human beings'.

[2] Functions words are regarded as expressing personal intentions, commands and desires

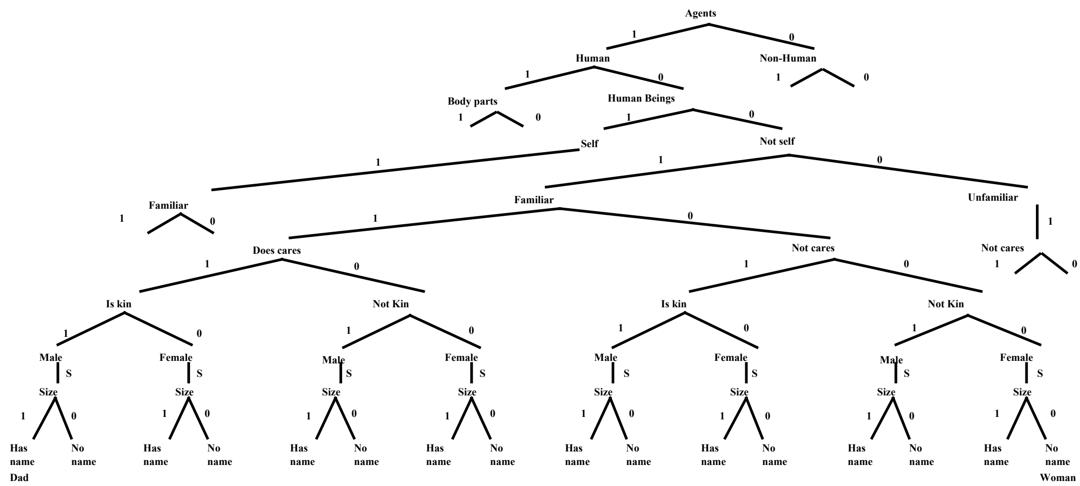

Figure 5: A cross-section of the individual feature tree for the category 'agents'. This figure shows the individual features for concepts identifying Human beings

Now that we have derived individual feature trees for Bloom's (and Anisfeld's) taxonomy, it is possible for us to represent and incorporate within our representation scheme the various children's concepts reported by Bloom. This is achieved by attaching concepts to the terminal nodes of the individual feature tree. The individual features of a particular concept, for instance 'dad', can then be obtained by translating the constituent individual features into binary digits:

Agents -> Human -> Human Beings -> Not self -> Familiar -> Does care -> Is Kin -> Male -> Size (Large) -> Has name -> 'Dad'

Individual features for 'Dad' = [1, 1,0, 0, 1, 1, 1, 1, 1, 1]

We have presented a synthesis of both Nelson's and Bloom's descriptions to devise a concept representation scheme that takes into account the so-called 'defining features' that define super-ordinate categories, and the 'individual features' which uniquely identify Bloom's concept. In this way, the semantic feature vector encodes two types of information: super-ordinate category information (defining features) and specific information (individual features). Of course, this is an open question in semantics and in philosophy. However, our reasons for attaching a feature vector to each concept comprising defining and individual features is purely pragmatic. Table 3 shows the semantic feature vectors for the concepts 'dad', 'mum' and 'dog' created from a synthesis of Nelson's semantic structure and the individual features derived by us from child language literature.

Concept Instance	**Superordinate category**	**Individual Features**	**Semantic Feature Vector**
dad	object - animate - people - specific	agents, human, human-beings, not self, familiar, does cares, is kin, male, large, has name	[**1,1,1,1,** 1, 1, 0, 0, 1, 1, 1, 1, 1, 1]
mum	object - animate - people - specific	agents, human, human-beings, not self, familiar, does cares, is kin, female, large, has name	[**1,1,1,1**, 1, 1, 0, 0, 1, 1, 1, 0, 1, 1]
dog	object - animate - animal - generic	agents, non-human, animal, is indoor, furry coat, unfamiliar, no distinct colour, has distinct, sound, medium, no name	[**1,1,0,0**, 1, 0, 1, 1, 1, 0, 0, 1 , 0,0]

Table 3: Semantic feature vectors for concepts - 'dad', 'mum', and 'dog'. The defining features are given in bold type face

5. Simulation of the development of 'concept memory'

In connectionist terms, children's concept memory, or the so-called 'semantic store' where the acquired conceptual knowledge is 'stored', can be characterised by (a) the concept representation scheme, (b) the organisation of stored concepts, (c) the means for learning new concepts, and (d) the mechanisms for retrieving stored concepts. We believe that children's concepts, comprising objects, people, places or events that the child comes in contact with, can be represented and categorised in terms of a number of 'semantic features'. The learning of new concepts can to a certain extent be regarded as an unsupervised process, whereby children appear to detect the salient 'semantic features' of an concept without any guidance. The storage of concepts is effected by categorising them on the basis of perceived semantic features.

Assuming that the child must take some initiative during concept development, we have simulated concept memory comprising 43 'concepts'. This was achieved by using a 121 unit Kohonen map (cf. Section 2) - that uses the unsupervised learning regime. The 43 concepts were selected from the range of concepts reported in child language literature (Bloom 1973) and each concept is represented by a 20-dimensional semantic feature vector.

At the start of the simulation of the development of the concept memory, the Kohonen map implementing the concept memory contains no *a priori* knowledge. During learning the set of semantic features, that are to be learnt as patterns of correlated features, provide an inductive basis for demarcating the 2-dimensional output layer of the Kohonen map into categories or areas of close concepts. However at the start of the simulation, if the semantic feature representations of the concepts to be learnt are mapped on this 'randomly initialised' Kohonen map, one may observe that potentially close concepts are mapped sparsely, indicating the absence of any prior categories (see later Figure 7a).

The simulation of the development of the concept memory is carried out in an iterative manner, such that in each iteration a different concept is presented to the concept memory. The repeated presentation of the concepts over a number of iterations is analogous to the child's increased appreciation and knowledge of the concept over a period of time, and perhaps for the child it is this frequent repetition of information which leads to its assimilation. Individual concepts are presented more than once in a random order to ensure that the 'learning' taking place is not biased in any way and does not reflect a predefined course of development.

Given that the concepts are learnt, i.e. represented by individual units in the Kohonen map, learning can be quantified in terms of two parameters - (i) activation level (ACT) of the desired concept's unit when retrieved[3] and (ii) the 'Euclidean Distance' (ED) between the desired concepts' unit and the most highly active unit. In fact, as learning progresses, the ED is minimised by the self-organisation mechanism inherent in Kohonen maps, whereas at the same time the activation level of the desired concept unit increases. A concept is deemed to be learnt when, upon presentation of its semantic feature vector, the activation level of its representative unit is the highest amongst all other units (approaching unity), and its ED is the lowest (close to zero). Learning a concept ensures that it is retrieved when its corresponding semantic feature vector is presented to the concept memory.

By way of describing this complex simulation involving 43 concepts, we discuss the learning profile of just four concepts: 'dog', 'juice', 'dad' and 'cow' out of the 43 concepts to be learnt. The learning period spanned 8000 iterations. To provide a learning profile (shown in Table 4), we noted the amount of learning achieved after intervals of 500 iterations by taking a snapshot of the evolving concept memory.

[3]Concept retrieval in a connectionist network involves the presentation of a semantic feature vector - input pattern to the concept memory. This results in all units acquiring some activation level. The unit with the highest activation level best represents the input pattern, hence the concept associated with this unit is considered retrieved. If adequate learning has been achieved, the concept retrieved corresponds to the input pattern, otherwise other similar concepts may be undesirably retrieved.

Iteration	'DOG'	'JUICE'	'DAD'	'COW'
Range	RU	RU	RU	RU
1 - 500	pig	--	dad	cow
501 - 1000	dog	juice	mum	horse
1001-1500	duck	juice	mum	horse
1501-2000	duck	juice	mum	horse
2001-2500	dog	juice	mum dad	cow horse
2501-3000	dog	--	dad mum	cow horse
3001-3500	dog	--	dad mum	cow horse
3501-4000	dog	cokie	dad mum	cow --
4001-8000	dog	juice	dad	cow

Table 4: Learning profile showing the development of concepts: 'juice', 'dad' and 'cow'. RU indicates the 'Retrieved Unit'.

Table 4 shows that at the very first iteration, the ED between the (random) weight vector of all the units and the input stimulus is computed. The unit that has the minimal distance to the stimulus is 'assigned' the stimulus label. Subsequent iterations involve the computation of the ED and the reassigning of concepts to the units. After 500 iterations, when the stimulus 'dog' was presented to the concept memory, it retrieved the concept 'pig' - the Kohonen map has not yet learnt to discriminate between a 'dog' and a 'pig' and can easily confuse the two. This 'confused' behaviour of the Kohonen map can be explained as follows: the semantic feature representations of both concepts - 'pig' and 'dog' share a number of features. The retrieval of the proximate concept 'pig' instead of the concept 'dog' clearly indicates that, at this stage, the Kohonen map has acquired an understanding of a category structure, i.e., the defining features have been learnt. However, the Kohonen map is still not able to discriminate among the individual features of the concepts 'dog' and 'pig' (both concepts belong to the same category) and therefore confuses the stimulus 'dog' with the close concept 'pig'. Figure 6 shows graphs for concept development, both in terms of activation level and ED.

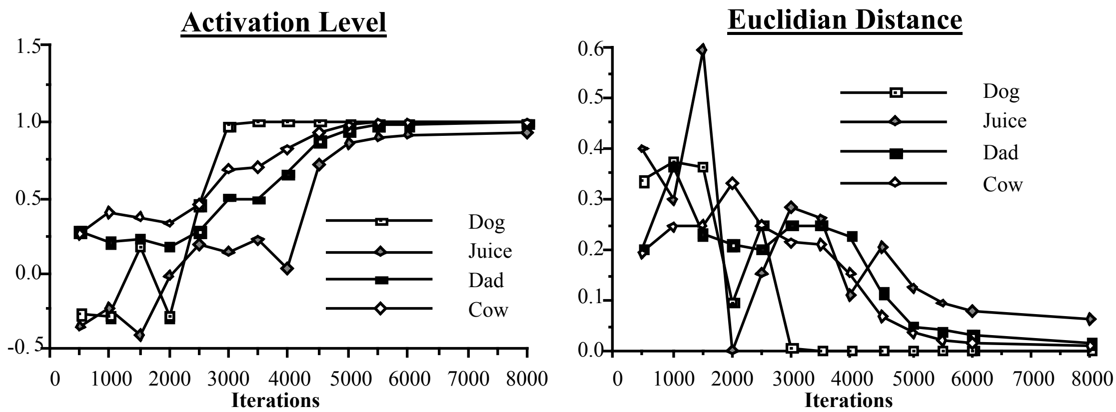

Figure 6: Learning profile in terms of both Activation Level and Euclidean Distance

At the end of 1000 iterations, the stimulus 'dog' retrieves the unit labelled 'dog', but the value of the ED is quite large (0.372) and the activation level is very low, in fact it is negative (-0.29): this retrieval may yet turn out to be a 'fluke'. This is confirmed at the end of 1500 and 2000 iterations; the Kohonen map now confuses the concept 'dog' with 'duck'. But after 2500 iterations, we see in Figure 6 a positive activation and a reduction of the ED in the learning profile for the concept 'dog'. Subsequent iterations do show that the network is becoming more 'stable' in its response to the stimulus 'dog': a doubling of the activation level between 2500 and 4000 iteration and a 200 fold reduction in the ED (see Figure 6). At iteration 4000, the criteria for adequate learning have been satisfied, i.e., the activation level has approached unity and the ED has decreased to zero (see Figure 6).

The learning profile for the other three concepts - 'juice', 'dad' and 'cow' follow a similar trend as noted in the development of the concept 'dog', such that the activation level starting from a low value increases towards unity and an initially high ED is reduced to zero. Note that for the concepts 'dad' and 'cow' during the iteration range 2000-4000 (shaded grey in Table 4) an interesting behaviour is observed. When presented with the semantic feature vector for a concept, say 'dad', two concepts are retrieved: the concept 'dad' and another close concept - 'mum'. This rather atypical behaviour predicates the fact that during this iteration range the Kohonen map is not able to differentiate between close concepts in a category. The retrieval of all the close concepts clearly indicates that at this stage the Kohonen map has learnt a category structure, i.e. defining features, and is exploiting this information when deciding what concepts are to be retrieved. However, the individual features of concepts need yet to be learnt. Figure 7b shows the organisation of the concept memory after a learning session of 8000 iterations, where each concept is represented by a unique unit. It is interesting to compare how (in Figure 7b) the concept memory has originated from the randomly initialised concept memory shown in Figure 7a.

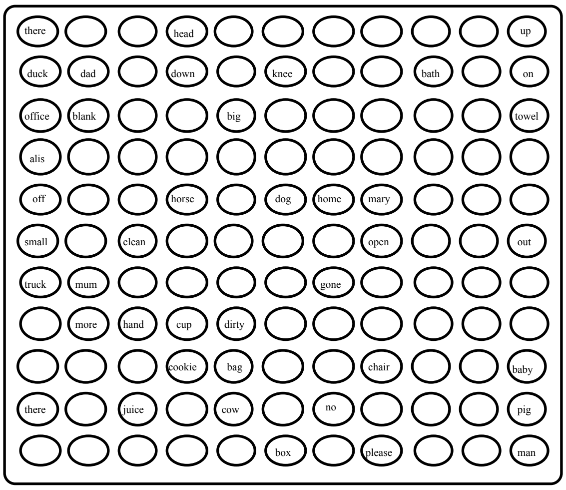

Figure 7a: Concept Memory before learning

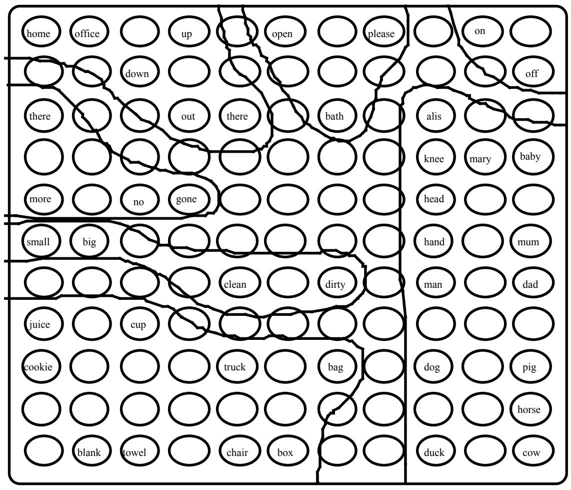

Figure 7b: Concept Memory after learning

5.1. A side-effect of the simulation: Emergence of a category structure

<u>'Automatic' Categorisation:</u> The organisation of the concept memory in Figure 7b reveals that concepts having close semantic feature representations are actually stored in proximity, thus forming a global organisation into conceptual regions or, more appropriately, 'categories' of concepts. It may be observed in Figure 7b that, the 'learnt' concept memory is divided into seven broad concept categories suggested by Bloom (1973) - *objects, agents, locations, attributes, prepositions, events* and *function words*. We have drawn the final lines in Figure 7b to emphasise this categorisation of concepts. Similar categorisation effects were reported in psychological experiments by Rip, Shoben and Smith (1973). It is interesting to note that during learning the connectionist network was not provided any category information nor an explicit definition of the semantic features and the possible relationships among them.

<u>Local organisation inside a category</u>: The same categorising principle which earlier formed the categories based on 'defining features' is again responsible for creating a local organisation or 'sub-category' of even closer concepts within a category. For instance, in Figure 7b the *agent* category includes the concepts *dad, mum, Mary,* and *man* that share a number of 'individual features'; hence these concepts are stored in proximity to each other, thus forming a sub-category, say 'humans'.

<u>Prototypical concept within a category or sub-category</u>: The category structure learnt by the concept memory (Kohonen map) is such that it is possible to extract prototypical information across many exemplars, while simultaneously storing idiosyncratic information about individual exemplars.

5.2. Addition of new concepts to a 'learnt' concept memory

Child theorists have speculated that the learning of a new concept is constrained by children's prior information about the environment. The categorisation of concepts helps in the learning of new concepts as a new concept can be perceived in terms of an existing concept, i.e., the features of a new concept are compared with the features of concepts in a particular category. For instance, the child may identify a new concept 'cat' in terms of a known and similar concept 'dog', because the new concept 'cat' has features such as 'animal', 'has tail', 'has furry coat', 'roams in the house', 'is pet', etc. which are common to the child's existing concept of a 'dog'.

In our connectionist concept memory addition of new concepts takes into account the prior existence of a taxonomy of earlier learnt concepts. We demonstrate this aspect by adding a new concept 'cat' to the previously learnt concept memory (shown in Figure 7).

It may be noted that the new concept 'cat' (shaded dark in Figure 8) is learnt and mapped in the immediate proximity of the concept 'dog' by the Kohonen map's self-organising learning mechanism. This indicates three things: (a) the connectionist learning mechanism is aware of the existence of a category structure, (b) the learning mechanism not only 'automatically' determined the category of the new concept but also figured out the sub-category to which it belonged, and (c) within the sub-category the concept 'cat' was placed next to the concept which bears greatest similarity to it (i.e. the concept 'dog').

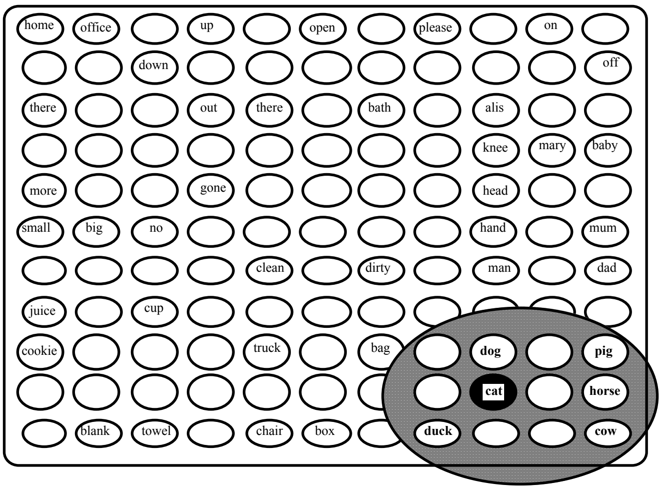

Figure 8: The concept memory with the newly added concept 'cat'. The new concept 'cat' is shown shaded dark. The light shaded area contains all *animal* concept and can be regarded as the area representing the sub-category 'animals' within the category 'agents'.

An explanation for the above behaviour is that the Kohonen map's self-organising learning mechanism earlier tuned the units in the neighbourhood of the concept 'dog' towards its semantic feature representation, thereby creating an area where the neighbouring units of the concept 'dog' have an internal representation that is close to it. Later, when the new concept 'cat' was presented to the Kohonen map to be learnt, it was mapped on to one of the neighbouring units of the concept 'dog' due to its similarity with the concept 'dog' .

The connectionist learning mechanism for adding new information has then the following characteristics: unsupervised learning, no re-organisation of the initial memory structure to accommodate new information, implementation of Piaget's notions of 'assimilation' and 'accommodation' (cf. Section 5.3), automatic and intelligent detection of the category of the new concept, and storage of new concept in proximity of similar concepts.

5.3. 'Indirect' simulation of Piaget's notions of assimilation and accommodation

We argued earlier that learning in ACCLAIM is based on Piaget's learning principles of assimilation and accommodation. Assimilation is evident during learning when a new concept partially activates other similar concepts, implying that the new concept is recognised in terms of previously stored concepts, where similarity is in terms of shared semantic features. The so-called accommodation process is analogous to the structural alterations (weight changes) made to the Kohonen map to 'accommodate' the new concept. Therefore, it can be concluded that the simulation of the development of concepts is in the background of Piagetian notions of assimilation and accommodation.

6. <u>A simulation of the production of two-word sentences</u>

We mentioned earlier that ACCLAIM is a synthesis of various connectionist modules (cf. Section 3), each simulating some aspect of child language development. When each module has learnt its designated knowledge, be it concepts, words, semantic relations or else, then two-word sentences can be produced by exploiting the knowledge learnt by each module. The production of child-like two-word sentences requires an interaction of the knowledge learnt by the various modules of ACCLAIM: concepts to be uttered are retrieved from the concept memory, and their corresponding lexical labels are retrieved from the word lexicon; the semantic relation between the two concepts is determined by the semantic relation network; and the hypothesis about the correct word order is evaluated by the word order hypothesis testing network. The ability to produce two-word sentences, which in fact is the final output of ACCLAIM, is a measure of the success of our connectionist simulation.

Children use (two-word) sentences to talk about themselves, their needs, beliefs, desires, state or otherwise, to express some aspect of their environment. Children's utterances can be regarded as a means to an end: the child communicates to achieve some desired goal. Whenever a child produces a sentence he/she always has an underlying 'intention' on which the utterance is based (Small, 1990: 133-5). One may argue that the child expresses his/her 'intention', together with related concepts, by using words strung together in a 'meaningful' manner to form a sentence. Based on the above assumption, we consider the child's two-word sentence to comprise two concepts: one concept representing the child's communicative 'intention', and another concept (corresponding to an external perceptual stimuli) representing some aspect related with the 'intention'. We briefly explain a simulation of the production of a two-word sentence which corresponds to a real situation (see Table 5) reported in Bloom's (1973: 235) data.

Situation	**Two-word Sentence**
(Allison holding out box to Mommy). What ? (Mommy opens box; giving it to Allison).	Mommy open

Table 5: A real-life situation concerning Allison talking to her mother (Bloom, 1973: 235)

Given the above situation, it can be inferred that Allison's (the child) communicative 'intention' is to request an *'Action'* (to open a box of cookies) to be accomplished by an *'Agent'* - 'Mommy' or 'mum'. Therefore, the two concepts which form the input to the simulation are: First concept: The child's communicative 'intention' represented by the concept 'open'. Note that the concept 'open' is a member of the category *action*. Second concept: The perceptual stimuli corresponding to the *agent* concept 'mum'. The simulation of the production of two-word sentences comprises the following three stages:

Stage 1: Retrieve the two input concepts and their corresponding words

The input for simulating the production of two-word sentences is two concepts - 'open' and 'mum'. Since the child's concepts are transformed into words which are then combined together to form a sentence, in the first step we retrieve the two concepts from the concept memory and their corresponding words from the word lexicon. First, the concept 'open' is retrieved by presenting to the concept memory a semantic feature vector corresponding to the concept 'open' (cf. Section 4). The word 'open' corresponding to the concept 'open' is retrieved by exploiting the naming connections (the concept lexicalisation network) established between the concept memory and word lexicon. In a similar manner, the second concept 'mum' is retrieved along with its corresponding word 'mum'.

Stage 2: Determine the semantic relation between the two concepts

In the next stage, the semantic relation between the two retrieved concepts needs to be determined. The knowledge of semantic relations has earlier been learnt by the semantic relation network. Note that a semantic relation comprises the 'first concept category' combined with the 'second concept category'. The input to the semantic relation network therefore consists of the concept categories of the two concepts - 'open' and 'mum'. Figure 9 shows the final state (activation level of various category units) of the semantic relation network, where the active category units are shaded with degrees of grey with respect to their activation level. According to this scheme the higher the activation level of a category unit the darker the shade of grey.

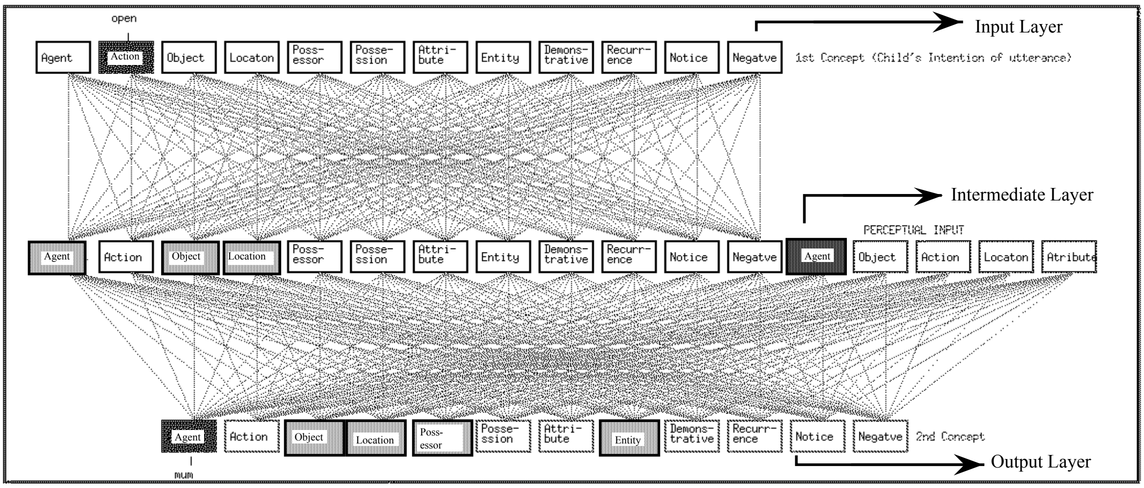

Figure 9: The final state of the semantic relation network, showing the activation level of the various category units in different layers. The 'Action' category unit in the input layer and the 'Agent' category unit in the output layer are the most highly active category units.

At the input layer, the first concept category 'Action', representing the input concept 'open', is presented to the semantic relation network resulting in the 'Action' category unit acquiring a high activation level (shown as a dark shade of grey in Figure 9). The high activation of the 'Action' category unit is spread across the connections to the category units in the intermediate layer. This results in all intermediate category units that may possibly have a semantic relation with the category 'Action' acquiring a high activation level. In this case, the category units 'Agent', 'Object' and 'Location' are activated with an equal activation level (as shown in Figure 9). Since no category unit in the intermediate layer is more active than the other active units, ACCLAIM is unable to determine the 'actual' second concept category which is semantically related to the first concept category 'Action'.

To determine the identity of the actual second concept category among the multiple candidate categories in the intermediate layer the perceptual input *mum* is presented to the semantic relation network. ACCLAIM deduces the concept category of *mum* to be 'Agent' (shown with a dark shade of grey in the perceptual input box in Figure

9) and regards it as corresponding to the second concept category. Next, activations are spread from the intermediate layer to the output layer. This results in the category unit 'Agent' in the output layer acquiring the highest activation level amongst all other category units in the output layer. This happens because the output layer's category unit 'Agent' receives combined activations from the 'Agent' unit in the intermediate layer and the perceptual input's 'Agent' unit. Therefore, we finally have a category unit - 'Agent' at the output layer which is highly active as compared to other active category units, and corresponds to the second concept category.

At the end of this stage, ACCLAIM has determined that the first concept category is 'Action' and the second concept category is 'Agent' and hence the semantic relation between the two concepts 'open' and 'mum' is **Action <--> Agent**.

Stage 3: Determine the correct word order
In the final stage ACCLAIM needs to determine the word order, i.e. how the two concepts 'open' and 'mum' are to be arranged in a sentence so as to reflect the word order observed in everyday adult language. The word order hypothesis testing network tests whether the correct word order should be (a) the first concept followed by the second concept or (b) the second concept followed by the first concept. Evaluation of either hypothesis is done by noting the error produced for each hypothesis. The hypothesis which produces the least error is considered to represent the correct word order. In this case, the correct word order is determined to be the 'second concept' followed by the 'first concept' (see Figure 10). ACCLAIM arranges the two words 'open' and 'mum', according to the determined word order to produce a two-word sentence - **'mum open'**.

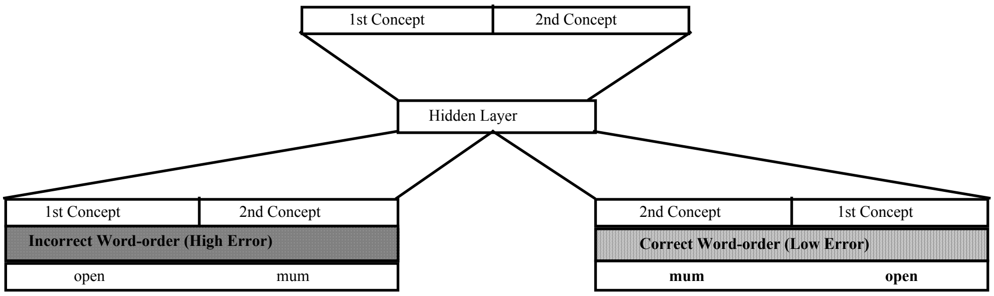

Figure 10: Evaluating word order hypothesis. The high error is shown by a darker shade of grey, whereas the low error has a lighter shade. The two-word sentence produced is 'there cookie'.

Table 6 gives a comparison of Allison's and ACCLAIM's response to some situations taken from Bloom's data (1973).

Situation	Allison's Response Two-word Sentence	Inferred Semantic Relation	ACCLAIM's response Two-word Sentence
(Mother pointing to chair.) What is this?	***that chair***	demonstrative + entity (Entity = object)	***that chair***
(Mother pours herself juice). (Allison picking up empty cup)	***more juice***	recurrence + object	***more juice***
(Mother pours juice; Allison drinks juice, looks into empty cup. Mother taking cup).	***gone juice***	negative + object	***gone juice***
(Allison reaching for cookie box in bag). (Allison takes out box of cookies)	***there cookie***	location + entity (Entity = object)	***cookie there***

Table 6: A comparison of Allison's and ACCLAIM's response i.e. the two-word sentences produced in the given situations

7. Conclusion

We have demonstrated how connectionist network architectures can be used to simulate different aspects of child language development. This was achieved through the use of archives of child language data collected by other researchers. These researchers have investigated aspects of language that are typical of children.

We have simulated different interactive facets of language development by choosing connectionist architectures that appear to be intuitively correct. Concept lexicalisation was simulated by the use of two self-organising networks, one for concept memory another for word lexicon, interconnected by a third Hebbian network. Semantic relations between concepts were learnt by our connectionist system in an unsupervised manner, and word order through supervised learning.

Our simulations of one-word and of two-word child language were based on a clear distinction between what the child him/herself makes of the perceptual features and phonetic input - unsupervised learning - and the role that the environment plays - supervised learning.

The emergence of 'connectionist network' architectures, architectures with a propensity for learning from being instructed or from experience, provides substantial opportunities for operationalising child language data and for testing child language theories.

References

Anisfeld, Moshe (1984)
*Language development from birth t*o three. London: Lawrence Erlbaum Associates

Bechtel, William. & Abrahamsen, Adele (1991).
Connectionism and the mind. Oxford: Basil Blackwell.

Bloom, Lois (1993)
Language acquisition and the power of expression. In Roitblat, H., Herman, L. and Nachtigall, P. (eds.), Language and communication. New Jersey, Lawrence Erlbaum Associates.

Bloom, Lois (1973)
One word at a time. Paris: Mounton.

Brown, Roger (1973)
A first language: The early stages. Cambridge, MA: Harvard University Press.

Caudill, M and Butler, C. (1992)
Understanding Neural Networks: Computer Explorations. Cambridge (Mass.): The MIT Press

Cromer, Richard (1991)
Language & thought in normal & handicapped children. Oxford: Basil Blackwell.

Hill, Jane (1983)
A computational model of language acquisition in the two-year-old. Cognition and Brain Theory, vol. 6, pp. 287-317.

Hinton, Geoffrey (1986)
Learning distributed representations of concepts. In Proceedings of the eighth annual Cognitive Science Society Conference, Amherst.

Kohonen, Trevo (1984)
Self-organisation and associative memory. Berlin: Springer-Verlag

Langley, P. (1982)
Language acquisition through error recovery. Cognition and Brain Theory, vol. 5, pp. 211-255.

Levine, David (1991)
Introduction to neural and cognitive modelling. Hillsdale: Lawrence Erlbaum Associates.

Markman, Ellen (1989)
Categorisation and naming in children: Problems of induction. Cambridge, MA: MIT press

MacWhinney, Brian (1987).
The Competition Model. In B. MacWhinney (eds.) Mechanisms of Language Acquisition, Hillsdale: Lawrence Erlbaum.

McClelland, John. (1989)

PDP: Implications for cognition and development In Morris, R. (eds.) Parallel distributed processing. Implications for psychology & neurobiology. Oxford: Clarendon Press.

McClelland, John (1988)
Connectionist Models and Psychological Evidence. Journal of Memory and Language, vol. 27, pp. 107-123.

Nelson, Katherine (1992)
Taxonomic knowledge: What kind and when?. Child Development, vol. 63, pp. 978-998.

Nelson, Katherine (1973a)
Structure and Strategy in Learning to Talk. Monographs of the Society for Research in Child Development. vol. 38.

Plunkett, Kim. & Marchman, V. (1989)
Pattern association in a back-propagation network: Implications for child language acquisition. Technical Report no. 8902, Centre for research in language, University of California, San Diego.

Rips, L. Shoben, E. & Smith, E. (1973)
Semantic distance and the verification of semantic relations. Journal of verbal learning and verbal behaviour, vol. 12, pp. 1 -20.

Rumelhart, David & McClelland, John (1986)
Parallel Distributed Processing, Vol I & II. Cambridge: MIT Press.

Selfridge, M. (1986)
A computer model of child language learning. Artificial Intelligence, vol. (29), pp. 171 - 216.

Seidenberg, M. & McClelland, John (1989)
A distributed developmental model of word recognition and naming. Psychological review, vol. 96(4), pp. 523 -568.

Shultz, Thomas (1991)
*Simulating stages of human cognitive development with connectionist model*s. In Birnbaum, L. & Collins, G. (Eds). Machine learning: Proceedings of the eighth international workshop, San Mateo, CA. Morgan Kaufman. pp. 105-9.

Siskind, J. (1990)
Acquiring core meaning of words, represented as Jackendoff-Style conceptual structures, from correlated streams of linguistic and non-linguistic input. In Proc. of the twenty-eighth annual meeting of the association for computational linguistics, University of Pittsburgh, Pittsburgh, Penn.

Small, Melinda (1990)
Cognitive development. San Diego: Harcourt Brace Jovanovich Publishers.

Smith, L. (1989)
A model of perceptual classification in children and adults. Psychological review, vol. 96, pp. 125-144.

Smith, L. (1984)
Young children's understanding of attributes and dimensions: A comparison of conceptual and linguistic measures. In Child development, vol. 55, pp. 363 - 380.