In Malaysian Journal of Computer Science, Vol. 10(2), 1997

USING NEURAL NETWORKS TO EXPLICATE HUMAN CATEGORY LEARNING: A SIMULATION OF CONCEPT LEARNING AND LEXICALISATION

Syed Sibte Raza Abidi
School of Computer Sciences
Universiti Sains Malaysia
11800 Penang
Malaysia
Tel.: 604-6573335
Fax: 604-6573335
email: sraza@cs.usm.my

ABSTRACT
Neural networks, provide a basis for studying various aspect of human learning and cognitive development in a seemingly 'psychologically plausible' manner. We present a 'hybrid' neural network architecture comprising two Kohonen maps interrelated by Hebbian connections to perform a neural network based simulation of the development of a 'concept memory', 'word lexicon' and 'concept lexicalisation' in an unsupervised learning environment using realistic psycholinguistic data. The results of the simulation demonstrate how neural networks, incorporating unsupervised learning mechanisms, can indeed simulate the learning of categories amongst children. The work demonstrates the efficacy of neural networks towards providing some insights into the elusive *mechanisms that lead to the emergence of human categories and an explication of inherent conceptual categories.*

Keywords: Neural Networks, Unsupervised Learning, Hybrid Architecture, Category Learning

1.0 INTRODUCTION

Neural Networks is a research discipline with the agenda to (a) understand the nature of human intelligence by simulating aspects of human behaviour; (b) incorporate 'human-like' intelligence within computer systems; and (c) provide a conceptualisation of the mind. These objectives are addressed by a highly parallel architecture comprising simple processing units that facilitates massive co-operative computation among these processing units. The processing units are provided with a variety of 'stimuli' and by communicating with each other they 'respond' in a manner that mimics aspects of human behaviour. The key notions in neural networks include *learning*: learning from being instructed by a 'tutor' in a supervised manner or learning on its own through observation and deduction in an unsupervised fashion.

The emergence of neural networks has provided scientists, both in computer science and psychology, an alternative framework for understanding the intricate and elusive tenets of human cognition; and a case can be made that this new class of information processing models construe cognition not as involving symbol manipulation, rather neural networks focus on causal processes that facilitate the excitation and inhibition of simple, highly interconnected processing units. Such a methodological position makes neural networks a prime candidate for modelling human cognition -- researchers recommend neural networks as a set of general principles, primitives, structures and approach that appears to provide a computatonal framework that is both psychologically and neurologically plausible [1]. In fact, the argument is extended further by arguing neural networks to be seen as a vehicle for operationalising cognitive development theories [2].

One important aspect of cogntive development is the emergence or learning of conceptual categories amongst humans, in particular developing children. Human learning of categories is a fairly indirect affair. Many categories that humans learn in real life are acquired in observational conditions in an autonomous learning environment that does not involve any feedback from a 'tutor'. In neural network terms this would have clear parallels with the unsupervised mode of learning as opposed to the supervised mode of learning based on feedback from a 'tutor'. If one believes that categories are a means of making sense of the environment then the argument goes that some aspects of learning about the environment should be unsupervised because humans, or more appropriately developing children, must invent their own categories for describing the environment as they perceive it.

The question we ask in this paper is how are categories learnt by humans? Indeed how one goes on to investigate how categories are learnt? There is no direct evidence. The only evidence that shows children have in fact learnt some categories is when they talk about groups of animals, groups of people, sets of toys, edible items, furniture and so forth. Child language researchers would argue that when children learn language, they appear to learn concepts first, or more precisely categories of concepts. Futhermore, the learnt concepts have been lexicalised - meaning that the child has not only learnt about concepts and categories of concepts but has also learnt to name them and articulate about them.

The lexicalisation of concepts, then, is a phenomenon which has to be simulated within the scope of neural networks in order to evaluate whether neural models can indeed simulate human learning and that the models demonstrate the emergence of categories during learning. This task would require the learning of concepts and associated words. We believe that such a study would provide students of human learning an operational framework where data and theory can be tested.

This paper describes neural network based simulations involvings the learning of concepts and the learning of associated words and furthermore discusses how one can use a 'hybrid' neural architecture to interrelate concepts and words, i.e., to simulate *lexicalisation* of concepts. In architectural terms, we present a 'hybrid' neural network architecture comprising two Kohonen maps and a Hebbian connection network. Note that Kohonen maps and Hebbian connections are both instances of unsupervised learning and we show how unsupervised learning algorithms can indeed learn categories by using realistic child language development. We believe that development of language involves continuous interaction of the child with the environment which in turn leads to the unsupervised 'invention' of categories. In neural network terms environmental influence, in terms of 'perceptual' and 'audio' (phonetic) stimuli, during learning is demonstrated by the adaptability of the 'plastic' structure of the neural networks to account for information received from the environment. In passing, we would like to note here that not much work has been undertaken on the learning of categories using unsupervised learning neural networks.

2.0 INTRODUCTION TO NEURAL NETWORKS

Neural networks attempt to mimic the neural structure of the brain *albeit* rather simplistically in that a neural network comprises a large number of computationally simple processing units. The processing units are highly interconnected through *plastic* connections. The 'plasticity' in the architecture of a neural network is introduced with the help of varying *connection weights* that can change over time and with experience. Basically, the connection weight determines the effect of the incoming input on the activation level of the unit. The configuration of the neural network dynamically adapts to the environment as a consequence of 'learning'. Put simply, learning in neural networks can be envisaged as the problem of finding a set of connection weights which allow the neural network to store experiential knowledge and to exploit it to simulate the desired behaviour. One can then argue that neural networks have a 'natural' propensity for storing experiential knowledge which is acquired and retained through 'training' or 'learning' as opposed to explicit programming.

Typical explanations of neural network learning 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" [2: pg. 270]

Neural networks learning algorithms are broadly classified into two main categories: supervised learning and unsupervised learning. Supervised learning algorithms require an input pattern along with a desired output pattern. The learning algorithm typically computes the difference between the desired output of the network to its actual output, i.e. an *error* value.. The computed error is then used to modify the interconnections between the units. Best exemplars of supervised learning are perceptrons and backpropagation networks. Unsupervised learning a lgorithms relate to the so-called 'self-organising' networks. Here, the neural 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. Kohonen maps [3] are amongst one of the best examples of this class of neural networks and are particularly useful for organising and categorising complex, multidimensional information.

3.0 A 'HYBRID' NEURAL NETWORK ARCHITECTURE

A psychologically plausible simulation of category learning involves the simulation of three distinct, yet highly interrelated, psychological activities – (i) the development of concepts, (ii) the learning of words and (iii) the lexicalisation of concepts – associating concepts with corresponding words. These activities can be further distinguished by the existence of a variety of input patterns, representation schemes, outputs and the underlying processing requirements.

To perform a realistic simulation of human category learning we propose a 'hybrid' neural architecture that synthesises three individual neural networks: (1) *Concept Memory*- characterising children's 'semantic store' where the acquired conceptual knowledge (i.e. concepts) is stored; (2) *Word Lexicon* - characterising children's 'phonological store' where words corresponding to concepts are stored; and (3) *Naming Connection Network* - storing associative relationships between concepts and their lexical labels, i.e. names. For our purpose, a 'hybrid' neural network integrates in a principled manner a number of neural networks, where each neural network simulates a particular psychological activity. The efficacy of hybrid neural network architecture originates from the architectural and functional synthesis of the neural networks and a co-operation between the constituent neural networks yields the overall objectives of the simulation.

The choice of appropriate neural networks for each activity, and more so how these neural networks are to be synthesised to form a hybrid architecture, is far from universal and formal. However, our choice of neural networks for simulating each activity is guided by psychological observations pertaining to the task. Furthermore, in order to maintain psychological plausibility we base our selection of the neural networks, from the wide range of available neural network classifications, architectures and learning algorithms, on the following criterion: (a) type of data to be learnt, (b) input data representation formalism, (c) explication of output and (d) learning strategy involved.

Psychological evidence suggests that all three activities that are to be simulated involve an unsupervised mode of learning. For that matter we have chosen an unsupervised learning neural network – the so-called Kohonen maps [3] for simulating the development of the concept memory and the learning of words. Kohonen maps employ a 'self-organising' algorithm for learning; in fact the efficacy of Kohonen maps is further extended by the fact that the learning algorithm segregates the input space into distinct regions or 'topological maps', where each region may contain similar patterns – the so-called automatic categorisation of patterns.

Development of naming connections is also to be simulated by using an unsupervised learning algorithm – Hebbian connections that are regarded to be the simplest algorithm for learning associations between two entities. Architecturally, the naming connection network connects <u>all</u> the output units in the concept memory with <u>all</u> the output units in the word lexicon. Appropriately weighted Hebbian connections, termed as 'naming connections', establish a relationship between a concept in the concept memory with its corresponding lexical label, i.e. word, in the word lexicon. These Hebbian connections are used to spread the activations from one Kohonen map to another such that a localised activity pattern in either Kohonen map will cause a corresponding localised activity pattern on the other Kohonen map, and this would be the basis of *concept lexicalisation*. Table 1 gives the architectural specifcations of the three neural networks to be used for the simulation with detailed description to follow in the forthcoming discussion.

Table 1: Architectural specifications of the hybrid neural network architecture

Activity	Input layer	Output Layer
Development of <u>Concept Memory</u> (*Kohonen Map*)	20 units	121 units
Development of <u>Word Lexicon</u> (*Kohonen Map*)	5 units	121 units
Development of <u>Naming Connections</u> (*Hebbian Connection Network*)	121 units	121 units

Now that we have specified the constituent neural networks we present a synthesis of these neural networks to realise our 'hybrid' neural network architecture (see Fig. 1) to carry out the simulations. A simulation model development framework for the synthesis of various neural networks to yield a conglomerate' neural network architecture has been proposed by Abidi [4, 5, 6, 7]. Here, the integration of the concept memory, word lexicon and the naming connection network is in line with the *modular* (or hybrid) architectrure approach proposed in detail by Abidi [4, 5, 6, 7].

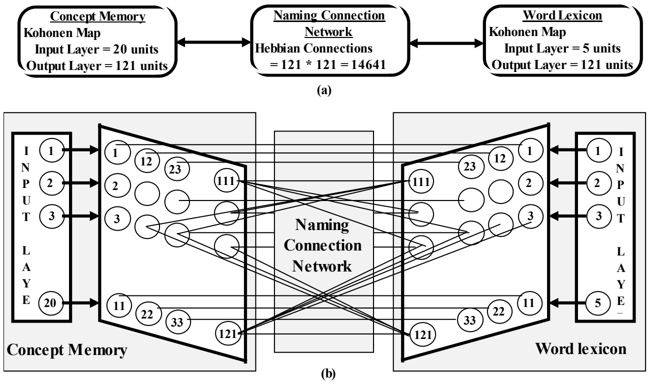


Fig. 1: The 'hybrid' simulation architecture:A synthesis of two Kohonen maps via Hebbian connections

4.0 A SIMULATION OF THE DEVELOPMENT OF THE 'CONCEPT MEMORY'

In neural network terms, children's concept memory 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. Assuming that the child must take some initiative during concept development, we regard the learning of new concepts 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. We have simulated the development of a 'concept memory', i.e., the learning of 42 'concepts' using a 121 unit Kohonen map. The 42 concepts being learnt were selected from the 50-60 concepts reported in child language literature [8].

Previous simulations of the development of concepts were conducted in a supervised learning environment [9, 10, 11]; a concept was learnt by repetitively associating its semantic feature based representation with the concept's lexical label. Categorisation was achieved by merely learning the label of the category to which it belonged. We have some disaggreements with this strategy to learn concepts as it seems more as a case of 'rote learning'. We elaborate below our simulation of the development of the concept memory together with the discussion on the emergence of conceptual categories.

4.1 A Neural Network Inspired Concept Representation Scheme

Represenation of natural conepts in a neural network formalism is of prime importance and demands a fine balance between psychological plausability and neural network pragmatism. To represent concepts in a neural network environment we have adopted the conventional 'semantic feature' based formalism which describes the similarities and differences between various concepts that leads to the definition of categories. Each concept in our representation scheme is represented by a 20-dimensional 'semantic feature vector' [12] 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 are based on an 'object-oriented' taxonomy suggested by Katherine Nelson [13]. Nelson's 'semantic structure' classifies or categorises 'objects' and 'non-objects' at a considerable level of detail, enabling us to determine the category of the object/non-object concept in consideration.

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. We argue that 'individual features' unique to a concept help discriminate one concept from other concepts having the same 'defining features'. For instance, children are believed to distinguish various objects by observing aspects such as 'size', 'shape', 'colour' and even, at times, their 'function'. For that matter, the individual features derive from a taxonomy of children's concepts suggested by Bloom [8], comprising concepts belonging to seven different categories: *objects, agents, events, states, locations, prepositions* and *'function words'*.

To conclude, our semantic feature vector encodes two types of information: super-ordinate category information (defining features) and specific information (individual features). Table 2 illustrates the semantic feature vectors for some exemplar concepts using in our simulation.

Table 2: Semantic feature vectors for concepts - 'dad', 'mum', and 'dog'. The defining features are given in bold type-face.

Concept Instance	Defining Features	Individual Features
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]	

At the start of the simulation of the development of the concept memory, the Kohonen map implementing the concept memory is initialised with random weight vectots. This ensures that the concept memory does not contains any *a priori* knowledge. This claim is validated by noting that potentially close concepts are mapped quite sparsely, indicating the absence of any prior categories. (see Fig. 2a).

Kohonen map's learning can be quantified in terms of two parameters - (i) activation level (ACT) of the desired concept's unit when retrieved 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 map learning algorithm, whereas at the same time the activation level of the desired concept's unit increases. A concept is deemed to be learnt when the activation level of its representative (or image) unit higher that all other units (approaching unity), and its ED is the lowest (close to zero).

4.2 Description of the Simulation

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 [14]. 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. Presentation of individual concepts in a random order ensures that the 'learning' taking place is not biased and does not reflect a predefined course of development.

In order to describe this complex simulation involving 42 concepts, we discuss the learning profile of just four concepts-'dog', 'juice', 'dad' and 'cow' out of the 42 concepts to be learnt. The learning period spanned 8000 iterations. To provide a learning profile we noted the amount of learning achieved after intervals of 500 iterations by taking a snapshot of the evolving concept memory. The learning profile of the concept memory is given in Table 3.

Table 3 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 amongst the individual features of the concepts 'dog' and 'pig' (since both concepts belong to the same category) and therefore confuses the stimulus 'dog' with the relatively close concept 'pig'.

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 negative (-0.29): this retrieval may yet turn out to be a 'fluke'. This is justified at the end of 1500 and 2000 iterations; the Kohonen map now confuses the concept 'dog' with 'duck'. But after 2500 iterations, one sees 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'. 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.

Table 3: Learning profile showing the development of concepts - 'juice', 'dad' and 'cow'. RU indicates the 'Retrieved Unit' in response to a specific concept

Iteration	Dog RU	Juice RU	‘Dad RU	Cow 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 --
4001-4500	dog	juice	dad	cow
4501-5000	dog	juice	dad	cow
5001-5500	dog	juice	dad	cow
5501-6000	dog	juice	dad	cow
6001-8000	dog	juice	dad	cow

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'. Note that for the concepts 'dad' and 'cow' during the iteration range 2000-4000 (shaded grey in Table 3) an interesting behaviour is observed. When presented with the semantic feature vector for the concept 'dad', two concepts are retrieved: the concept 'dad' and another close concept - 'mum'. This rather atypical behaviour predicates the fact that 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 and is exploiting this information when deciding what concepts are to be retrieved.

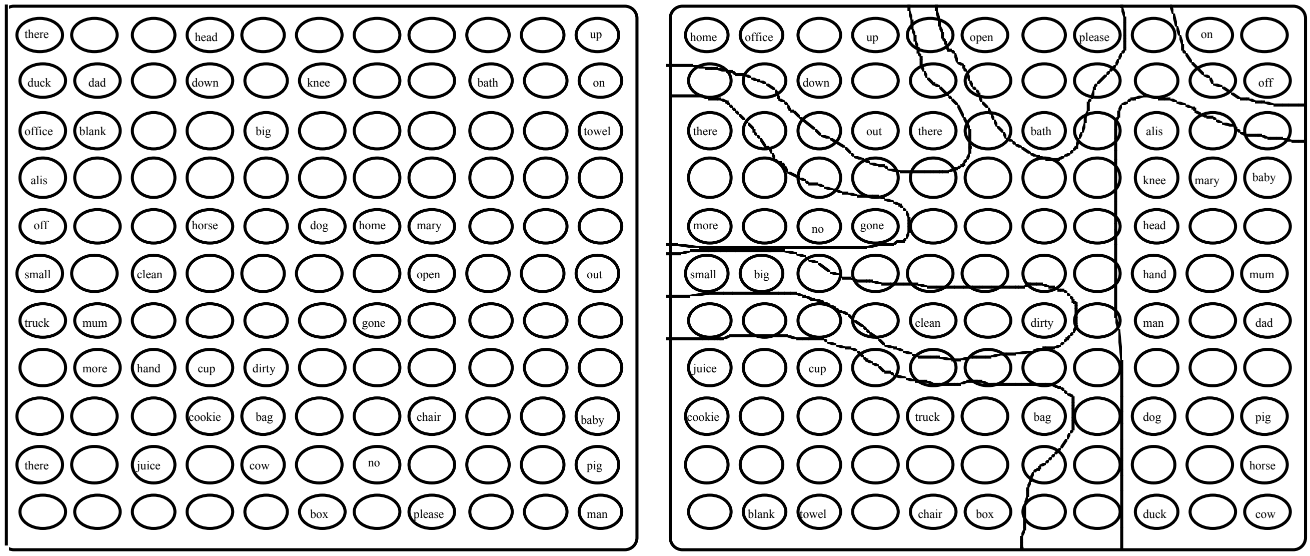
Fig. 2b shows the organisation of the concept memory after a learning session of 8000 iterations, where each concept is represented by a unique unit. It is interesting to compare how the concept memory has originated from the randomly initialised concept memory, shown in Fig. 2a.

5.0 A SIMULATION OF THE DEVELOPMENT OF THE 'WORD LEXICON'

One significant manifestation of the development of language amongst children is their ability to comprehend and produce spoken language. One can model this aspect of language development by arguing that children can analyse acoustic input in terms of its constituent phonemes. The ability to 'spot' words in continuous speech can be compared with the development of the so-called 'similarity neighbourhoods'. -- “a set of words that differ from a given target by a phoneme substitution, addition or deletion” [15: pg. 207]. The concept of similarity neighbourhood relates to the fact that similar sounding 'words' would be represented in a cluster or 'category'. For instance, the similarity neighbourhood for the word *pit* would include the words *bit, pot, pig, spit, and it,* amongst others.

From a neural network standpoint, then, one can argue that, given phonetic input to a Kohonen map (the so-called word lexicon), the output from it construes to be a set of words corresponding to different phonetic inputs. Also, the organisation of these words in the word lexicon predicates a discrimination of phonetic information leading to a 'similarity neighbourhood' that seem analogous to the categorisation of the word lexicon, which results as a consequence of the temporal organisation of phonetic information.

For the development of the word lexicon, we 'train' a 121 unit Kohonen map to initially learn and then to recognise 'words' given their phonetic representation. The phonetic representation of each word is taken from the Oxford Advanced Learner's Dictionary. To represent words, we have devised an encoding scheme which assigns each phoneme a numerical value within the range of 0 -1.



(a) (b)

Fig.2: (a) The concept memory before learning. (b) The concept memory after learning. The concept memory is divided into seven broad concept categories - *objects, agents, locations, attributes, prepositions, events and function words*

The phonemic representation for a word, i.e. its 'phonemic feature vector', is formed by concatenating the encoded value of its constituent phonemes in a vector notation. For illustration purposes, the phonemic feature vector for some words is given in Table 4.

Table 4: Phonemic representation of words in terms of a phonemic feature vector

Word	**Phonemic Feature Vector**
dog	[0.45, 0.60, 0.65, 0.0, 0.0]
bag	[0.25, 0.40, 0.65, 0.0, 0.0]
pig	[0.15, 0.20, 0.65, 0.0, 0.0]
dad	[0.45, 0.40, 0.45, 0.0, 0.0]

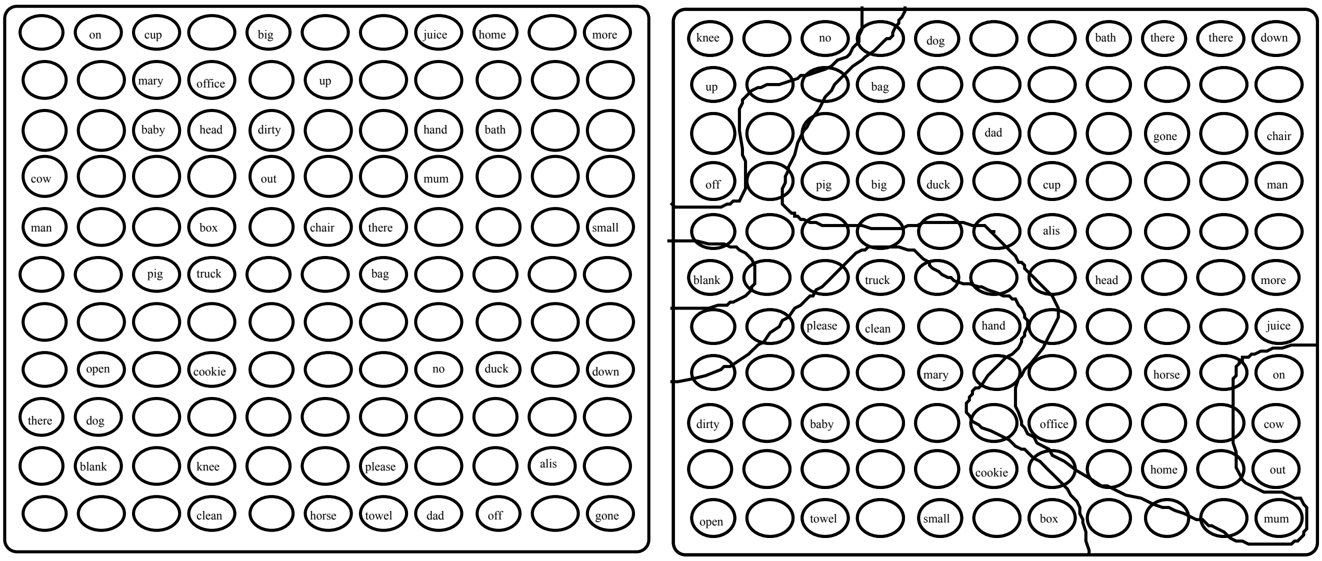
The simulation of the developing word lexicon is performed in a similar manner to that of the developing concept memory. Starting with a random Kohonen map, phonemic feature vectors of words are presented in a random order. The learning profile of the word lexicon follows a similar trend as that of the concept memory, and again the criteria is the activation level approaching unity and the ED being reduced to zero.

Fig. 3a shows the initially random word lexicon, whereas Fig. 3b shows the word lexicon after the learning session. In Fig. 3b it can be seen how the word lexicon has evolved from a random organisation of words to an ordered organisation that reflects categorisation of words on the basis of the length of the phonemic feature vectors. Again, like we did for the learnt concept memory, we have marked regions of the Kohonen map that store words of similar phonetic lengths. It is these regions that resemble the 'categories' or 'similarity neighbourhoods' argued by researchers [15]. Note that the 'learnt' word lexicon clearly discriminates words on the basis of their phonetic content, and also within categories similar sounding words are stored in proximity, for instance note that the similar sounding words - 'bag', 'dog', 'pig', 'big', 'dad' and 'duck' are stored close to each other.

6.0 A SIMULATION OF THE DEVELOPMENT OF THE 'NAMING CONNECTIONS' – *CONCEPT LEXICALISATION*

In child language literature, lexicalisation or 'naming' of concepts is regarded as the mapping of children's linguistic knowledge on to their conceptual knowledge [16,17]. Lexicalisation of a concept can loosely be regarded as either learning by instruction or learning from examples. We believe that there are at least two 'ostensive naming' situations that can be simulated by neural networks conducting 'unsupervised learning'. The first situation relates to the assignment of a word to a 'known concept' where the child has a concept of an object or event but lacks the appropriate word to express it. The second situation relates to the assignment of a word to a 'novel concept': The child hears a novel word referring to an novel object or event, then the child relates the novel word to the new concept.

In both the above situations the child need to identify the category of the input concept and retrieve it from the concept memory. Also during word perception the demand on the child is to analyse the phonetic constituents of the word and retrieve the correct word, if present, from the word lexicon. The lexicalisation of concepts then is a viable simulation to further explicate and operationalise 'learnt' categories within the concept memory and word lexicon.



(a) (b)

Fig. 3: (a) Word Lexicon before Learning. (b) Word Lexicon after learning. Words are arranged into four categories on the basis of their phonetic length, i.e. words constituting 2, 3, 4, and 5 phonemes

6.1. Impetus for the Simulation Architecture

We have simulated concept lexicalisation as the development of an association between a lexical label (word) with the corresponding concept. In a neural network parlance such an association would be achieved by learning associative *naming connections* between a concept unit in the concept memory with the corresponding word unit in the word lexicon. Simulation of concept lexicalisation involves each unit in the concept memory to be connected, with varying connection strengths, to all units in the word lexicon and vice versa. This establishes a many-many relationship between concept units and word units, as shown in Fig. 4.

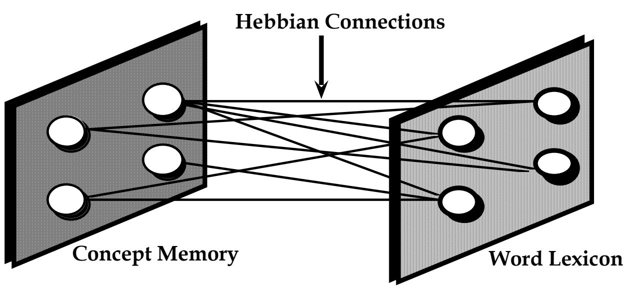


Fig. 4: Naming connections between the Concept Memory and the Word Lexicon.

The naming connection network, simulating concept lexicalisation, employs Hebbian Learning which provides a simple mechanism for associating two units by weighted connections, where the strength of the connection is based on the activation levels of the two connecting units - the greater the collective activation level of the two units the stronger the 'Hebbian conection' between them. The naming connection network, comprising two layers, employs the output layers of both the concept memory and the word lexicon which are then connected by 14641 (121 * 121) Hebbian connections of varying strengths. The 'learnt' Hebbian connections provide a medium to tranmit the activation level of the output units of the concept memory to the output units of the word lexicon and vice versa. The strength of the Hebbian connetion determines the amount of activation received by the recipient output unit – the greater the connection weight between two units the magnified the activation level of the sending unit would be when it is received by the recipient unit.

6.2. Simulation Scheme for Developing Naming Connections

Two types of stimuli were used in this simulation: (1) a perceptual stimuli, i.e. a 20-dimensional semantic feature vector representing a concept, and (2) a phonemic stimuli, i.e. a 5-dimensional phonemic feature vector representing the corresponding word. The entire training set comprises Bloom's 42 concepts and words learnt earlier.

The development of naming connections can be simulated in a developmental manner by simultaneously presenting a concept (the perceptual stimuli) to the concept memory and the corresponding word (phonemic stimuli) to the word lexicon. Presentation of the respective stimuli to each Kohonen map results in a group of units to become highly active due to the information retrieval mechanisms employed by Kohonen maps. Naming connections can then be established between these highly active units, in each Kohonen map, based on the Hebbian learning algorithm.

Here we discuss the lexicalisation of the concept 'dog'. Consider an exemplar situation for concept lexicalisation: an adult points towards a 'dog' and utters the sentence *'That is a dog'*, thus both the verbal and perceptual stimuli corresponding to 'dog' are presented to the learner.almost at the same instance. The information retrieval mechanism of the Kohonen maps ensures that the presentation of the perceptual stimuli to the concept memory forms a localised pattern of activity around the learnt 'dog' concept unit. In this scenerio similar concepts are more activated than less similar concepts. In a similar manner, the presentation of the verbal stimuli 'dog' to the word lexicon results in the learnt word unit 'dog' acquiring the highest activation level. At this stage, we apply the Hebbian learning algorithm to establish inter-map naming connections amongst all units in both Kohonen maps. The strength of the Hebbian connection established is proportional to the current activation of two connecting units. Therefore, a strong connection is established between the highly active concept and word units, i.e. the 'dog' concept and word units.

Concept lexicalisation is carried out in an iterative manner, where in each iteration a concept-word pair is presented to the naming connection network and learning involves slight increments to the strength of the Hebbian connections between the concept and word units. In this way, over a period of several iterations strong naming connections are established between concepts and their corresponding lexical labels (words). A concept is deemed to be lexicalised when a 'perceptual' stimuli representing a concept is presented at the concept memory and in response the lexical label - 'word' unit corresponding to the concept is highly active in the word lexicon.

Table 5 presents the learning profile of the lexicalisation of four concepts - 'dog', 'cow', 'juice' and 'dad' which are represented in the 'learnt' word lexicon by units 76, 88, 55 and 86, respectively (Recall that each unit in Kohonen map has been assigned a number in the range 1 - 121). For instance, at iteration 500 the concept 'dog' is associated with unit 2 in the word lexicon. This turns out to be an incorrect association since the actual word unit representing 'dog' is 76. During subsequent iterations the neural network is again incorrectly associating the concept 'dog', first with word unit 36, and then later with word unit 114. It is only after 6000 iterations that the neural network has learnt to lexicalise the concept 'dog', as now the concept 'dog' is associated with word unit 76, which represents the word 'dog'. The learning profile for the other three concepts show a similar trend where first incorrect associations are established between concept and word units in the concept memory and word lexicon, respectively. However with increased experience the correct associations are eventually 'learnt'.

It may be noted (see Table 5) that during the lexicalisation of a particular concept, say 'dog', not only the concept 'dog' is associated with the word 'dog' but also other similar category members are associated with the word 'dog' though with a less strong connection. This ensures that a strong naming connection is established between the close category members and less strong naming connections exist among other not so close category members.

Table 5: Learning profile for *Concept Lexicalisation.* Finally, the words dog, cow, juice and dad are represented by units numbered 76, 88, 55 and 86, respectively

Iteration Range	**Dog RU**	**Cow RU**	**Juice RU**	**Dad RU**
1 - 500	2	36	70	17
501 - 1000	36	**88**	56	17
1001-1500	36	**88**	56	17
1501-2000	36	**88**	91	17
2001-2500	114	**88**	91	17
2501-3000	114	**88**	56	17
3001-3500	114	**88**	56	17
3501-4000	114	**88**	91	17
4001-6000	114	**88**	91	119
6001-6500	**76**	**88**	91	119
6501-7000	**76**	**88**	91	119
7001-7500	**76**	**88**	**55**	**86**
7501-8000	**76**	**88**	**55**	**86**

7.0 EXPLICATING THE EVIDENCES OF CATEGORY LEARNING – THE SIMULATION RESULTS

The aim of the paper is to demonstate the emergence of human conceptual categories. Our assumption, which is psychologically motivated, is that it is best to investigate the emergence of categories at the onset of concept development as usually whilst learning new concepts humans distinguish and discrimate various concepts;

grouping and subgrouping of dynamicalaly concepts are made and periodically refined with time and with increased appreciation of existing concepts, thereby realising categories of similar concepts. In the absence of any direct means to investiugate how categories are learnt we exploit our three 'learnt' neural networks – the concept memory, word lexicon and naming connection network to explicate the subtle evidences implicit category learning during the processes of concept development, word acquisition and concept lexicalisation. We now explicate both direct and indirect (by exploiting the three neural networks) evidences of an underlying category structure that has been implicity learnt by the three neural networks during the simulations.

7.1 'Automatic' Categorisation of Concepts

The organisation of the concept units in the concept memory reveals that concepts that have close semantic feature representations are actually stored in proximity, thus forming a global organisation into conceptual regions or, more appropriately, 'categories' of concepts (see Fig. 2b). Effectively, self-organisation in Kohonen maps demarcates the possible input space into hierarchical sub-areas which are then mapped on to the two-dimensional Kohonen map. In Fig. 2b we have marked the Kohonen map to explicate the emergent categories of concepts. The reader may note that the right side of the concept memory accommodates concepts of the category 'agent', whilst 'object' concepts are stored in the bottom left corner and similarly the 'location' category occupies the top left area of the concept memory.

Note that the semantic feature representation of each concept is based on a hierarchical structure: 'defining features' (containing category information) and 'individual features' (distinguishing individual concepts). Whilst learning the concepts, the Kohonen map exploited the category information and collected concepts with similar 'defining features'. These semantically close concepts were then stored in proximity to each other, resulting in clusters of concepts that resemble 'categories'. In this way the Kohonen map not only learnt the concepts, but also simulated an 'automatic categorisation' of the concepts.

It is interesting to note that during learning the neural network was not provided any category information nor explicit definition of the semantic features and the possible relationships among them. Nonetheless, the Kohonen map itself deduced the similarity among the 'defining features' of various concepts and 'automatically' created clusters or categories of close concepts.

7.2 Local Organisation Inside a Category - Presence of *Sub-Categories*

The same categorising principle which earlier formed global categories based on 'defining features' is again responsible for creating a local organisation or 'sub-categories' of even closer concepts within a category. This local organisation is a manifestation of the similarities among the 'individual features' of various concepts belonging to the same category (see Fig. 2b). Put simply, the Kohonen map's learning algorithm analyses the finer distinctions in the semantic feature vector of concepts belonging to the same category and then organises close concepts in proximity. For instance, in Fig. 2b the *agent* category includes 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'. Also, within the same *agent* category, concepts for animals such as *dog*, *pig, cow* and *horse* are in proximity to each other, thus resembling another sub-category - 'animals'.

Recall that a concept's semantic feature vector encodes an implicit hierarchy; the 'defining features' determine the broad category and the 'individual features' distinguishes concepts within categories. It is interesting to note that the Kohonen map, whilst learning the concepts, was able to detect this implicit hierarchy in the feature representation of the concepts, thus instigating two categorisation activities: the 'defining features' were used to determine a broad category structure, whereas based, on the 'individual features', concepts belonging to the same category were locally categorised.

7.3 Indirect Evidence of the Existence of Categories - *Concept Generalisation*

Ward and Vela [18] have reported that the manner in which children generalise from a novel or partially visible category exemplar to other members of the category is influenced by children's prior knowledge of previously learnt categories.

To investigate the presence of categories within the concept memory we tested the generalisation capabilities of the 'learnt' concept memory. This was achieved by presenting the concept memory with (a) an incomplete representation of a learnt concept and (b) a novel concept. For case (a) we presented an incomplete semantic feature representation of the concept 'dog'. In response the Kohonen map completed the partial representation and correctly retrieved the concept 'dog'. For case (b) we presented a representation of a novel concept - 'cat'. Again the 'learnt' Kohonen map determined the possible category of the novel concept, which is 'agents', and subsequently generalised the novel 'cat' concept to the closest learnt concept 'dog' in the 'animal' sub-category.

It may be noted that, much as what Ward & Vela have suggested, during generalisation the concept memory first determined the appropriate conceptual category to which the novel or partially represented concept may belong. Then, from the candidate conceptual category one concept that was most similar to the novel concept was selected.

7.4 Indirect Evidence of the Existence of Categories - *Addition of New Concepts*

Child theorists have speculated that the categorisation of concepts helps in the learning of new concepts as the new concept can be perceived in terms of an existing concept. For instance, the child may identify a new concept 'cat' in terms of a known and similar concept 'dog', in that the new concept 'cat' shares features such as 'animal', 'has tail', 'has furry coat', 'roams in the house', 'is pet', etc. with the child's existing concept of a 'dog'.

Our neural network based concept memory verifies the existence of such a behaviour, as is illustrated when attempting to add a new concept 'cat' to the previously learnt concept memory (shown in Fig. 5). It may be noted that the new concept 'cat' (shaded dark in Fig. 5) is learnt and mapped (in the areas corresponding to the category *agent* and the sub-category *animals*) in the immediate proximity of the concept 'dog'. This indicates three things: (a) the learning mechanism is aware of the existence of an implicit category structure underlying the organisation of the concept memory, (b) the learning mechanism not only 'automatically' determined the category of the new concept but also determined 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 resemblance to it, i.e. the concept 'dog'.

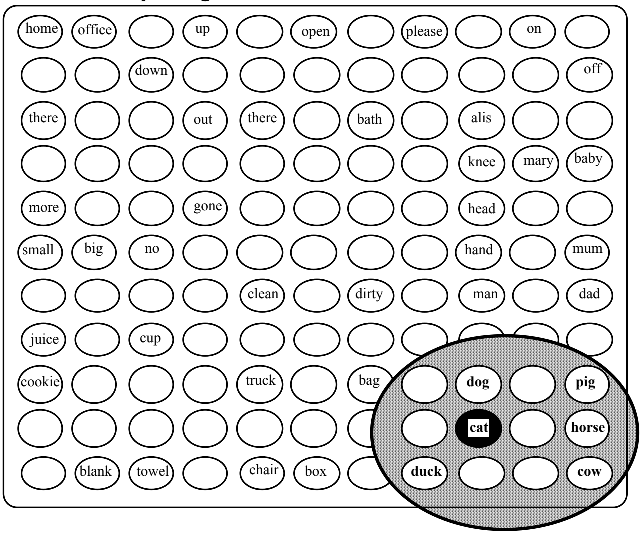


Fig. 5: The concept memory with the newly added concept 'cat'. The shaded area represents the sub-category 'animals' within the category 'agents'.

7.5 Indirect Evidence of the Existence of Categories – *Retrieval of a Concept-Word Pair*

Again, an indirect evidence of the presence of categories within the learnt concept memory and the word lexicon is available by simulating the retrieval of a word in response to a given concept. Here, the naming connection network is used to demonstrate *concept lexicalisation*.

Retrieval of a 'word' therefore involves an interaction among three neural networks - concept memory, naming connection network and word lexicon. The retrieval of a word in response to a concept is simulated by presenting a concept to the concept memory. This initiates the spreading of the activation of the concept units through the 'naming connections' to the word units. If a strong naming connection exists between a concept unit and its corresponding word unit, then the presentation of the concept to the concept memory enables the corresponding word unit in the word lexicon to acquire the highest activation level amongst all other word units.

We now demonstrate the retrieval of the word 'dad' when given the concept 'dad' to the learnt concept memory. To begin with, the presentation of semantic feature vector for the concept 'dad' is presented at the input layer of the concept memory. This brings into relief the information retrieval mechanism of Kohonen maps - the learnt concept unit 'dad' acquires the highest activation level and is deemed as being retrieved in response to the input (shown in Fig. 6). Next the naming connection network is used to retrieve the lexical label of the retrieved concept 'dad'. By employing the spreading activation mechanism the activation level of all active concept units is spread through the naming connections to the word-lexicon. This flow of activation results in the emergence of localised patterns of activations on the word-lexicon, such that word units that are strongly connected with the highly active concept units acquire a high activation level. In this case, the word unit 'dad' acquires the highest activation level and is deemed as being retrieved (shown in Fig. 7) in response to the concept 'dad'.

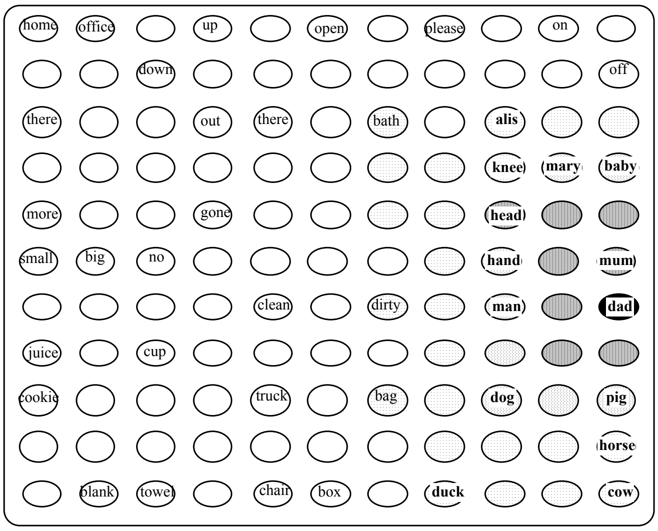


Fig. 6: State of the concept memory when presented the concept 'dad'. The degree of activation level is depicted by darker shades of grey. Concept unit 'dad' has the highest activation level.

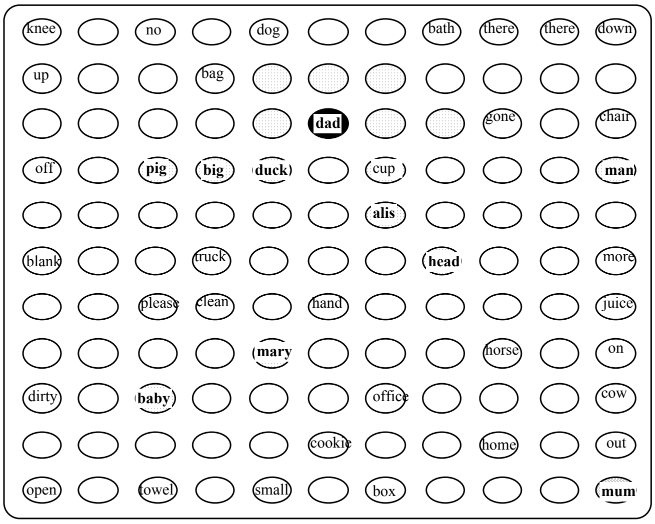


Fig. 7: State of the word lexicon after activations are spread from the concept memory. Word unit 'dad' is retrieved as it has the highest activation level.

In Fig. 6, it may be observed that the presentation of the concept ‘dad’s perceptual stimuli to the concept memory has resulted in concepts belonging to the category ‘agents’ to acquire higher activation levels as compared to other units in different categories, thereby suggesting the overall selection of the 'agent' category. It may therefore be argued that during the retrieval of the concept first the broad category was selected and subsequently the selection was narrowed down to one category member that best represented the perceptual stimuli, i.e. the concept ‘dad’.

In Fig. 7, it may be observed that apart from the highly active word unit 'dad' the word unit 'mum' is the next most highly active unit. This again explicates the network’s knowledge of an implicit category structure; the neural network has deduced that the concepts 'dad' and 'mum' are very similar to each other, and this conclusion is validated by the high activation level of words corresponding to concepts belonging to the ‘agent’ category.

We argue that this concept-word retrieval simulation not only demonstrates the information retrieval mechanisms inherent in Kohonen maps but also validates the efficacy of the Hebbian connections implemented in the naming connection network, and in turn proves to be a good means of explicating the category information learnt by the neural networks - the concept memory and the word lexicon.

8.0 CONCLUDING REMARKS

We have demonstrated that *neural networks* provide a basis for investigating human category learning. Our simulations showed effects of category learning during the development of concepts, associated words and the lexicalisation of concepts. The emergent categories are interpreted in terms of the neural networks partitioning or discriminating the input stimuli in an unsupervised learning environment. We have shown that such self-organising neural networks may have some parallel with human (category) learning.

From a neural network standpoint we have demonstrated the efficacy of a 'hybrid' neural architecture for simulating aspects of human behaviour. Recall that the neural networks were subjected to unsupervised training, a training regime that has empathy with the developmental paradigm of language development. Furthermore, the connections between the two networks was established fairly successfully, through what appears to be a training regime based on neo-Hebbian 'laws', rooted in the behaviouristic paradigm. The fact that a number of researchers in neurobiology, developmental psychology and linguistics are interested in neural networks and neural simulations leads us to believe that we have made a contribution towards some questions related to the understanding of human behaviour.

REFERENCES

[1] D. Rumelhart and J. McClelland, P*arallel Distributed Processing*, *Vol I & II.* Cambridge:MIT Press, 1986.

[2] W. Bechtel and A. Abrahamsen, *Neural Networks and the Mind*. Oxford:Basil Blackwell, 1991.

[3] T. Kohonen, *Self-organisation and Associative Memory*. Springer-Verlag, 1984.

[4] S. S. R. Abidi, “Neural Networks and Child Language Development: Towards a ‘*Conglomerate’* Neural Network Simulation Architecture*”* in *International Conference on Neural Information Processing (ICONIP’96), Hong Kong, 1996.*

[5] S. S. R. Abidi, “Integrating Supervised & Unsupervised Learning Strategies For Simulating Child Language Development” in *World Congress on Neural Networks (WCNN’96), San Diego, 1996.*

[6] S. S. R. Abidi, “A Neural Network Simulation of Child Language Development at the One-Word Stage” in *IASTED International Conference on Modelling, Simulation and Optimization, Gold Coast, 1996.*

[7] S. S. R. Abidi & K. Ahmad, “Conglomerate Neural Network Architectures: The Way Ahead for Simulating Early language Development”. *Journal of Information Science and Engineering,* Vol. 13, 1997, pp. 235-266.

[8] L. Bloom, *One Word at a Time*. Paris:Mounton, 1973.

[9] J. Clapper and G. Bower, “Category Invention in Unsupervised Learning”. *Journal Of Experimental Psychology: Learning, Memory And Cognition*, Vol. 20, 1994, pp. 442 - 460.

[10] G. Hinton, “Learning Distributed Representations of Concepts” in *Eight Annual Cognitive Science Society Conference, Amherst, 1986.*

[11] D. Medin, “Concepts and Conceptual Structure”. *American Psychologist*, Vol. 44, 1989, pp. 1469-81.

[12] K. Nelson, “Some Evidence for Cognitive Primacy of Categorisation and its Functional Basis”. *Merrill-Palmer Quarterly*, Vol. 19, 1973, pp. 21-39.
[13] K. Nelson, “Structure and Strategy in Learning to Talk”, *Monographs of the Society for Research in Child Development*, Vol. 38, 1973.
[14] S. S. R. Abidi & K. Ahmad, “Child Language Development: A Connectionist Simulation of the Evolving Concept Memory”. M. Aldridge (Ed.) *Child language*. Clevedon: Multilingual Matters Ltd, 1996.
[15] J. Luce & P. Luce, “Similarity Neighbourhoods of Words in Young Children’s Lexicons”. *Journal of Child Language*, Vol. 17, 1990.
[16] S. Levine & S. Carey, “Up Front: The Acquisition of a Concept and a Word”. *Journal of Child Language,* Vol. 9, 1982, pp. 645 - 657.
[17] M. Callanan, “How Parents Label Objects for Young Children: The Role of Input in the Acquisition of Category Hierarchies”. *Child Development,* Vol. 56, 1985, pp. 508 - 523.
[18] T. Ward & M. Vela, “What Makes a Vibble a Vibble?” *Child Development*, Vol. 60, 1989, pp. 215-224.

BIOGRAPHY

Syed Sibte Raza Abidi has a Ph.D. degree in Computer Science from Surrey Univ. (UK) and Master’s degree in Computer Engg. from Miami Univ. (USA). His research interests include Neural Networks, Telemedicine, Medical Informatics, Data Mining & Expert Systems. He has worked on ESPRIT (European Commission) projects and is currently working on various IRPA projects. His research efforts have been widely published in International conferences and journals. He is leading the *Intelligent Systems Research Group* at USM - Computer Sciences.