In Proc. of the International Conference on Neural Information Processing (ICONIP'96), Hong Kong, 1996

Neural Networks and Child Language Development: Towards a *'Conglomerate'* Neural Network Simulation Architecture

Syed Sibte Raza Abidi
School of Computer Sciences
Universiti Sains Malaysia, Penang, MALAYSIA
E-Mail: sraza@cs.usm.my

Abstract

Neural networks provide a basis for studying child language development in that such networks emphasise *learning*. We report a simulation of some key aspects of child language development during infancy. We argue that in order to simulate the uniquely human language learning, it is important to use a 'conglomerate' neural network architecture that integrates the collective strengths of a variety of neural networks in some principled fashion to take into account the diverse nature of inputs to and outputs from a child learning language. We present such a 'conglomerate' neural network architecture - ACCLAIM that integrates both supervised and unsupervised learning algorithms, to simulate the learning of *concepts*, *words*, *conceptual* and *semantic relations* and simple *word-order* rules, thus mimicking the production of child-like *one-word* and *two-word* language. The simulations carried out are 'language informed' as realistic child language data has been used for training the neural networks.

1. Introduction

Neural network community is keenly interested in simulating human learning, and indeed language learning provides an interesting framework to build sophisticated information systems of the future. Language is generally learnt in a 'naturalistic' setting; the setting is very noisy and the input to and output from a child does not always obeys a predetermined sequence and order. There is a premium on the child correcting his or her own errors. Furthermore, child language development appears to be evolutionary, dynamic and involves an interaction of multiple tasks such as motor co-ordination, concept development, categorisation, perception, biological growth and social influences. In this regard, neural networks offers mechanisms, such as adaptive learning, generalisation, self-organisation, feature extraction, and pattern-recognition that appear to have direct relevance towards a computational simulation of child language development.

In this paper we present a 'conglomerate' (or 'modular') neural network based information processing model - ACCLAIM (A Connectionist Child LAnguage development and Imitation Model). ACCLAIM is a large-scale model that comprises a variety of neural networks communicating with each other in a systematic manner to simulate aspects of child language development within the age group of 9-24 months. Architecturally, ACCLAIM aims to achieve a degree of psychological plausibility: language development aspects which can be construed to be innate development have been simulated by using unsupervised learning regimes, and environmentally-determined aspects of language development have been simulated by using supervised learning regimes. The combined learning potential of the various neural networks constituting ACCLAIM are exploited to simulate how *concepts* are acquired and lexicalised; how *words* represented as phonemic features are learnt; how *conceptual relations* are learnt and expressed as *one-word utterances*. Furthermore, to mark a transition from one-word to two-word language, we have simulated how children learn *semantic relations* between conceptual categories and also how children learn *word-order* rules. Finally, we demonstrate how all this learnt knowledge is used to produce child-like *one-word utterances* and *two-word sentences*. The simulations performed using ACCLAIM are 'language informed' such that the data used in 'training' the neural networks was derived from the archives compiled by prominent child language researchers.

2. A psycholinguistic model of child language development

We believe that a systematic 'language informed' simulation of child language development need to be based on a psycholinguistic framework derived from prominent child theories and should use realistic child language data. Our psycholinguistic framework for simulating child language is originally due to the eminent child psychologist and theorist Jean Piaget. Interpretations of Piaget's psychological notions in a neural network terminology have been suggested by McClelland (1989) and Bechtal & Abrahamsen (1991). Furthermore, various researchers including Shawley & Schopman (1990), Shultz (1991), Levine (1991) and Seidenberg (1993) have emphasised the role neural networks can play in simulating high-level cognitive tasks.

Consider the following model of child language development: (a) Cognitive development of a child takes place in 'stages' and is made possible through an interaction between two processes - *assimilation* and *accommodation*; (b) The child is involved in an on-going *conceptualisation* process, perceiving the environment in terms of a set semantic features to form new concepts and to recognise known concepts (Nelson, 1973); (c) The child's lexical growth is predicated by the child's ability to analyse phonemic information received from adult language, leading to the storage of words in terms of their phonemic constituents; (d) 'Ostensive' naming of concepts establishes a relationship between the child's concepts and words, such that, words are verbal manifestations of the child's conceptual knowledge; (e) The child's initial *one-word utterances* reflect their awareness about 'conceptual relations', such as *recurrence*, *disappearance* and so on (Bloom, 1973); (f) The transition from one-word to two-word sentences is marked by the child's ability to plan to manipulate the meaning of individual words in terms of 'semantic relations' which they learn and express by structuring underlying conceptual categories (Brown, 1973); (g) The child learns the underlying word order of the adult language (Brown, 1973); (h) the child's two-word sentences reflect the child's *intention* to communicate certain semantic relations pertaining to self or events happening in the environment (Brown, 1973).

From the above model of child language development it is clear that the child is an active information processor and his/her language development is seemingly due to a subtle interplay of inborn capacities, psychological makeup and environmental input. Hence, the above-mentioned 'processes' can be further distinguished by demarcating the environmental considerations from what can be regarded as the 'innate' ability of the brain to learn language. Put simply, our conjecture is that aspects of child language development that can be construed to be innate can be simulated by using 'unsupervised learning' algorithms, whereas the environmentally determined aspects of language development can be simulated by using 'supervised learning' algorithms. Hence, since the child appears to employ a variety of learning mechanisms during language development, a plausible approach to simulate language development would be to include in the simulation model all available learning algorithms that have some parallels with the child's overall learning strategy.

3. A framework for developing *conglomerate* neural network architectures

The discussion of the above mentioned psychological processes involved in child language development implies the existence of a variety of neural networks, each simulating a particular process. This brings into relief the need for a 'conglomerate' (or 'modular') neural network architecture: an architecture that integrates in some principled fashion both supervised and unsupervised learning algorithms, thus exploiting the collective strengths of a variety of neural networks to provide a more 'realistic' simulation. Thus, in a 'conglomerate' neural network architecture simulating language development, both the effects of the environment and that of self-learning or 'innate development' can be distinguished; the former can be simulated through supervised learning networks and the latter through unsupervised networks.

Development of conglomerate neural network architectures, in simple terms, requires the 'mixing and matching' of the relative strengths of a variety of neural networks. We propose a framework for developing conglomerate neural network architectures that distinguishes candidate neural networks on the basis of their intrinsic characterstics such as learning mechanisms, input/output representation schemes, environmental considerations and so on (Abidi, 1994). Our framework mainly emphasises (i) psychological and neurobiological distinctions between various neural networks when selecting neural networks to simulate certain tasks; (ii) architectural specifications - determining the number of layers, number of units in a layer, activation update functions and learning parameters; (iii) a plausible connectivity scheme by which various neural networks can efficiently communicate with each other; and (iv) variety of training strategies, including (a) one neural network learning its training data independently; (b) two or more neural networks learning their specified training data simultaneously; and (c) a co-operative training strategy where one or more neural networks transform the training data to a representation scheme that is interpretable by the principal neural network being trained.

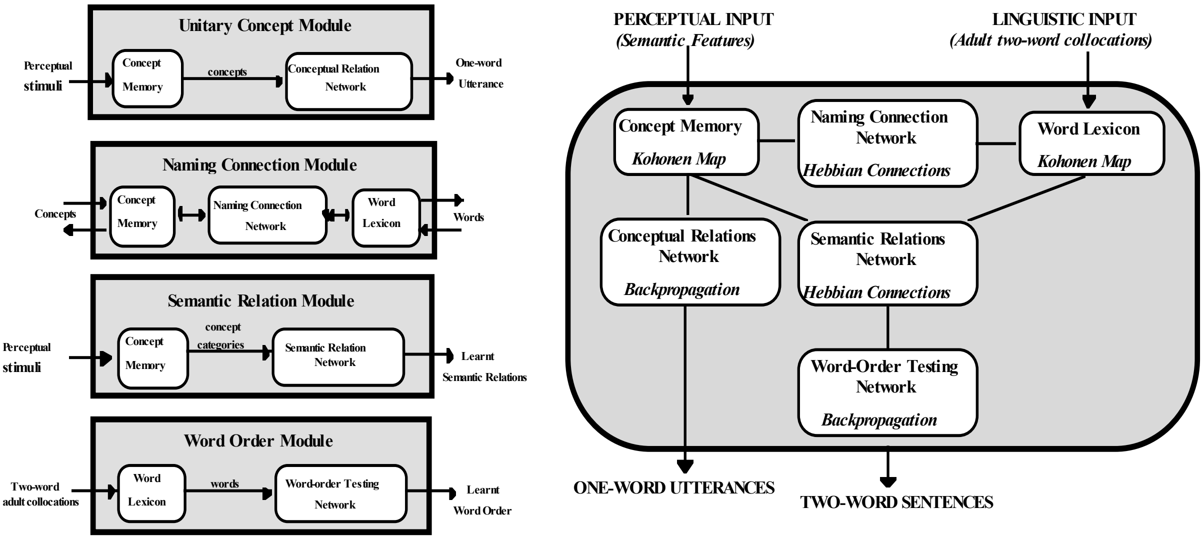
Table 1 lists the various neural networks that are used to implement the above mentioned model of child language development, together with a specification of the typical input and output for each process that may be involved in child language development.

Psychological Process	Typical Input	Typical Output	NN Simulating The Process	The NN's Specifications
Development of a *concept memory*	A vector comprising semantic features representing concepts	The storage of children's concepts and their categorisation	***Concept Memory***	**Kohonen Map** IP = 20 Units OP = 121 Units
Development of a *word lexicon*	Phonemic representations of words - Phonemic Feature Vectors	The storage of children's words and their categorisation	***Word Lexicon***	**Kohonen Map** IP = 5 Units OP = 121 Units
***Ostensive naming* of concepts**	Concepts (Semantic feature vectors) & words (Phonemic representations)	An association between children's concepts and the corresponding words - the *naming* of concepts.	***Naming Connections Network***	**Hebbian Connections** IP = OP = 121 Units
Learning *conceptual relations*	Conceptual relations, perceptual entities and functional words	An association between a 12 conceptual relations with 25 functional words (one-word utterances).	***Conceptual Relations Network***	**BP Network** IP = 15 units HI = 5 units OP = 25 units
Learning *semantic relations*	Two-word adult collocations and perceptual stimuli represented as conceptual categories.	Leant semantic relations (associative connections) among 12 concept categories.	***Semantic Relations Network***	**Hebbian Connections** IP = 12 units INT = 17 units OP = 12 units
Learning *word-order*	Auditory stimuli (Two-word adult collocations)	Learnt word-order rules based on adult language. Used to produce two-word sentences.	***Word-Order Testing Network***	**BP Network** IP = 12 units HI = 4 units OP = 12 units

Table 1: The various neural networks implementing the above-mentioned model of child language development. The table legend is IP = Input Layer, OP = Output Layer, HI = Hidden Layer, INT = Intermediate Layer

4. Towards the architecture of ACCLAIM

Language development is a complex activity and it seems improbable to perform a realistic simulation of language development using just a single neural network. We therefore propose a '*conglomerate*' (or 'modular') approach for simulating language development, whereby individual neural networks are configured in a meaningful manner to realise a 'neural network module'. Within a neural network module the individual neural networks retain their identity and merely interact with each other to provide a more powerful and elaborate response. Each neural network module can then be dedicated to simulate a particular aspect of child language development. For instance, a *concept naming module*, synthesising three neural networks - *concept memory*, *word lexicon* and *naming connections network* can simulate the ostensive naming of concepts. Using the individual neural networks (shown in table 1) as building blocks, we have developed four different modules, where each module can simulate a different aspect of child language development (figure 1a). Extending the modularity approach further, the various modules are then synthesised, based on the tenets of our psycholinguistic model, to realise the unified architecture of a 'conglomerate' language development model - ACCLAIM (figure 1b). A systematic interaction amongst all the constituent neural networks of ACCLAIM not only simulates child language development, but also produces child-like *one-word* and *two-word sentences*.



(a) **(b)**

Figure 1: (a) Four neural network modules each comprising more than one neural network and simulating some aspect of child language development; (b) The modular architecture of ACCLAIM - an integration of various neural networks each simulating an aspect of child language development

Indeed, a conglomerate approach for developing complex neural network architectures has certain advantages, for instance (i) it allows knowledge learnt by one neural network to be utilised in more than one module, for instance the concepts learnt and stored in the concept memory are used by three different modules - the unitary concept module, concept naming module and the semantic relation module (see figure 1a); (b) within a module one or more neural networks can be used to transform the representation of the training data to a representation that is understood by another neural network for both learning and information retrieval purposes; (c) the results of a simulation incorporating just one module can be used by other modules to perform their respective simulations; and (d) at a deeper level, each module again can be envisaged as an independent neural network model, capable of simulating a psycholinguistic process on its own.

5. A neural network based simulation of child language development

In ACCLAIM all the simulations were carried out in a 'developmental' manner, starting with no *a priori* information (a randomly connected organisation of processing units) the neural networks were repeatedly presented with a set of 'training patterns' which the networks gradually *learnt* over a period of iterations. This repeated presentation of the training patterns is analogous of the child's increased appreciation of interesting information, and perhaps it is this frequent repitition of information that leads to its assimilation. Table 2 summarises our simulation strategy both at the one-word and two-word stage of language development.

Simulation Task	Learning	Neural Networks Used
One-word language stage		
Concept development, storage and categorisation	Unsupervised	Concept memory
Concept generalisation, neologisms and novel concepts	Unsupervised	Concept memory
Word acquisition, storage and development	Unsupervised	Word lexicon
Learning conceptual relations - production of one-word utterances, generalisation to novel situations	Supervised	Conceptual relations network
Concept lexicalisation - development of naming connections	Unsupervised	Naming connection network
Concept and lexical retrieval	Unsupervised	Concept memory + Naming connection network + Word lexicon
Two-word language stage		
Learning semantic relations and determining the semantic relation between two conceptual categories	Unsupervised	Semantic relation network (involving Concept memory + Naming connection network + Word lexicon)
Learning word-order leading to the production of two-word sentences	Supervised	Word-order testing network (involving Concept memory + Naming connection network + Word lexicon + Semantic relations network)

Table 2: List of simulations characterising key aspects of child language development

Below we briefly describe some of the above-mentioned simulations:

Learning Concepts *(Concept Memory)*: Simulation involved the learning of 42 concepts (taken from Bloom, 1973) represented by a 20-dimensional semantic feature vector, comprising the so-called 'defining features' determining the concept's category and 'individual features' distinguishing the category members. The learnt concept memory exhibited clusters of close concepts or 'concept categories', thus implying an *automatic categorisation* of learnt concepts into categories and sub-categories (within categories).
Learning Words *(Word Lexicon)*: Simulation involved the learning of 42 words corresponding to the learnt concepts. The words were represented in terms of their phonetic components by a 5-dimensional phonetic feature vector. The learnt word lexicon predicated a discrimination of phonetic information, i.e. the development of the so-called 'similarity neighbourhoods' of similar sounding words.
Concept Lexicalisation *(Naming Connection Module)*: Simulation involved the creation of (bi-directional) 'naming connections' between each concept in the concept memory with its corresponding lexical label in the word lexicon. Thus, an activated concept/word would spread its activations acroos the 'learnt' naming connections leading to the retrieval of the corresponding word/concept.
Learning Conceptual Relations *(Unitary Concept Module)*: Simulation involved the learning of a mapping between a set of 12 *conceptual relations* and 3 *perceivable entities* (people, objects and events) to a set of 25 *words* (Bloom, 1973). The learnt conceptual relations network accepts a *conceptual relation* as input and as output produces the corresponding *one-word utterance* that best reflects the child's 'intention'. Also, the conceptual relation network can generalise to produce appropriate responses in novel situations.
Learning Semantic Relations *(Semantic Relations Module)*: Simulation involved the development of associative connections between 12 concept categories, implying that certain concept categories are 'semantically related' (Brown, 1973). Given a concept category its semantic relation with all other categories can be determined by spreading its high activation across the network to other connected concept category units.
Learning Word-order *(Word-Order Module)*: Simulation involved the learning of the inherent word-order in adult language, i.e. learning how to arrange two words to form a sentence. Given two words, the learnt word order testing network would combine two words in the proper order to yield a child-like *two-word sentence*.

6. Input transformation: Significance of the 'conglomerate' approach

Our conglomerate approach facilitates a co-operation amongst a number of neural networks for transforming the input stimuli from one representation to another representation. For instance, the input stimuli for the simulation of learning semantic relations is adult language (or simply adult two-word collocations), however what is actually needed for learning semantic relations is the conceptual category information of the two concepts that comprise the adult two-word collocation. This situation calls for the transformation of an adult two-word collocation to the corresponding conceptual category information, in a manner that may have some relevance to the processing of the child. By way of our conglomerate approach, the knowledge acquired by various neural networks can be exploited to transform an input stimuli to the desired representation. For learning semantic relations the *concept memory*, *word lexicon* and the *concept lexicalisation network* are employed according to the following scheme: learning semantic relations begins with the presentation of an adult-two word collocation to the word lexicon, which results in the retrieval of the corresponding two words. The naming connections are then exploited to retrieve the two corresponding concepts. Finally, the category information of the two retrieved concepts, embedded in their semantic feature representations in terms of the defining features, is extracted for the learning of semantic relations. Similarly, word-order learning requires the transformation of an two-word collocation to a semantic relation and this is achieved by an interaction between the *concept memory, word lexicon, naming connection* network and *semantic relation network* (see figure 2).

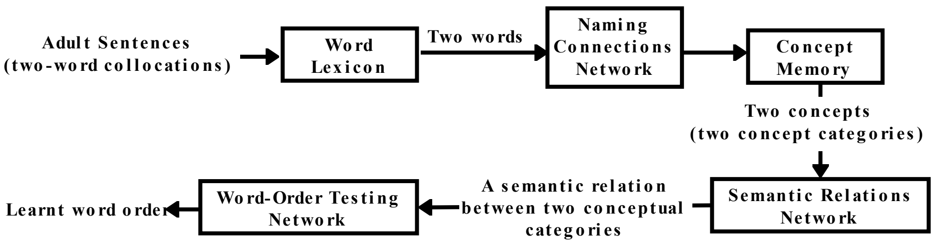


Figure 2: The input transformation procedure for simulating the learning of word-order

Indeed, one could have simulated both the learning of semantic relations and word-order by directly presenting category information to the concerned neural network, however our argument is that such a scheme would not hold much psychological plausibility as compared to our simulation scheme which brings into relief the underlying mechanisms that may be involved in the child's learning of semantic relations and word-order.

7. The simulation results: Production of *one-word* and *two-word language*

The learnt performance of ACCLAIM is best elicited by demonstrating the ability of ACCLAIM to produce child-like one-word and two-word language (as these two simulations involve an interaction between all the previously 'trained' neural networks) and quantifying the sentences produced against Bloom's (1973) data which reports children's utterances with details of the situation in which the utterance was made. Table 3 and 4 present a sample of the one-word and two-word sentences produced by ACCLAIM, respectively.

Simulation of one-word utterances production utilises the *unitary concept module*. The input stimuli comprising a conceptual relation (representing the underlying 'intention of the child') is presented to the concept memory. In response, the conceptual relations network produces a *one-word utterance* that best represents the child's intention. Simulation of two word sentence production involves an interaction between 5 neural networks, namely the *concept memory, word lexicon, naming connection network, semantic relations network* and the *word-order testing network*. This simulation, involving 3 stages, initiates with the presentation of an input stimuli which comprises two concepts - (1) the child's communicative 'intention', and (ii) a perceptual stimuli: (a) retrieval of two concepts from the concept memory and the corresponding words from the word lexicon; (b) determining the semantic relations between the categories of the two concepts; and (c) evaluation of a word-order hypothesis - determining the correct order in which the two words (corresponding to the concepts) are to be arranged. The two words when combined in the correct order yield a child-like *two-word sentence*.

Real-Life Situation (taken from Bloom, 1973)	Child's 1-word Utterance	ACCLAIM's Input Pattern	ACCLAIM's 1-word Utterance
(Mother pointing to chair) What is this?	***chair***	Pointing + Object (chair)	***obj name (chair)*** (output unit 17)
(Allison holding picture to photographer's assistant, off camera) Where's the girl?	***girl***		
(Allison turns picture over so she can't see the girl)	***no***	Non-existence + People	***no*** (output unit 7)
(Allison turning it back to picture side)	***there***	Existence + People	***there*** (output unit 5)

Table 3: A sample of the 1-word utterances produced by ACCLAIM compared with those produced by the child -Allison.

Real-Life Situation (taken from Bloom, 1973)	Child's 2-word Sentence	Semantic Relation	ACCLAIM's 2-word Sentence
(Mother pointing to chair.) What is this?	***that chair***	demonstrative + 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***

Table 4: A sample of the 2-word sentences produced by ACCLAIM compared with those produced by the child -Allison

8. Conclusions

We have attempted a simulation of some aspects of aspects of child language development during the key developmental periods, c.9-24 months. We have demonstrated how neural networks can be used to both operationalise extant child language corpora and systematise psycholinguistic theories. Furthermore, the results of our simulations indicate a degree of similarity between the learnt behaviours of the neural networks with the kind of behaviours exhibited by children whilst learning language. From a neural network point of view we have demonstrated the efficacy of a conglomerate neural network architecture for simulating high-level cognitive activities. The architecture of ACCLAIM and the resultant processing capabilities achieved, should be an indicator as to how functionally and structurally divergent neural networks when synthesised together in a meaningful manner, i.e. based on a psycholinguistic model, can simulate a high-level cognitive activity such as child language development. We believe that the simulations carried out by ACCLAIM has made a contribution to the on-going research regarding the role of neural networks in simulating human learning.

References

Abidi, S.S.R. (1994)
A Connectionist Simulation: Towards a Model of Child Language Development.. Doctoral Thesis. University of Surrey, Guildford, England.

Abidi, S.S.R. & K. Ahmad (1994)
Unsupervised Category Learning in a Hybrid Connectionist Architecture. Proc. of SEARCC '94 computer conference, Karachi.

Bechtel, W. & Abrahamsen, A. (1991)
Connectionism and the Mind. Oxford: Basil Blackwell.

Bloom, L. (1973)
One Word At A Time. Paris: Mouton.

Brown, R. (1973)
A First Language: The Early Stages. Cambridge, MA: Harvard University Press.

Levine, D. (1991)
Introduction to Neural and Cognitive Modelling. Hillsdale: Lawrence Earlbaum Associates.

McClelland, J. (1989)
PDP: Implications for Cognition and Development. R. Morris (Ed.) Parallel Distributed Processing: Implications for Psychology and Neurobiology. Oxford: Clarendon Press.

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

Shawley, A & Schopman, J. (1990)
Connectionism and the Foundations of Cognitive Science. Cognitive Systems, Vol. 2(4), pp. 373-382.

Seidenberg, M. (1993)
Connectionist Models and Cognitive Theory. Psychological Science, Vol. 4, pp. 228-235.

Shultz, T. (1991)
Simulating Stages of Human Cognitive Development with Connectionis Models In L. Birnbaum & G. Collins (Eds.) Machine Learning: Proc. of the Eighth International Workshop, San Mateo, CA. Morgan Kaufman.