Project no. 034362

ACORNS

Acquisition of COmmunication and RecogNition Skills

Instrument: STREP
Thematic Priority: IST/FET

D3.1 Memory Architectures

Due date of deliverable: 2007-11-30
Actual submission date: 2007-11-15

Start date of project: 2006-12-01 Duration: 36 Months

Organisation name of lead contractor for this deliverable:

Revision: 0.1

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

VERSION DETAILS	
Version:	0.3
Date:	20 Dec 2007
Status:	Final

CONTRIBUTOR(S) to DELIVERABLE	
Partner	***Name***
USFD	Mark Elshaw
	Roger K Moore
	Guillaume Aimetti

DOCUMENT HISTORY			
Version	***Date***	***Responsible***	***Description***
0.1	01/09/07	Mark Elshaw	writing
0.1	25/09/07	Roger K. Moore	editing
0.2	15/10/07	Mark Elshaw	Diagrams editing
0.3	20/12/07	Mark Elshaw	Incorporate comments of reviewers

DELIVERABLE REVIEW			
Version	***Date***	***Reviewed by***	***Conclusion****
0.2	17/12/2007	Michael Klein	Reviewed and changes included
0.2	17/12/2007	Lou Boves	Reviewed and comment considered

Table of Contents

1 Introduction

This ACORNS workpackage report considers the memory architecture required for the other workpackages in the project. It reviews memory models that have been developed in the past for various associated research fields. The intention is to provide a wide set of examples of these types of memory models, that are to be considered for inclusion in the ACORNS memory architecture. From this study a preliminary memory model is outlined that will have the capacity to perform pattern storage, discovery and retrieval within an overall architecture based on the memory-prediction model of Hawkins and Blakeslee (2004). In addition to the examination of the prediction-memory model there is the consideration of various memory components for inclusion in the ACORNS memory models such as attention, working memory, episodic memory, sensor/motor grounding systems, reinforcement learning as well as how these might feed into the memory architecture for the project.

2 Memory-Prediction model

The memory architecture developed as part of the ACORNS project in workpackage 3 is based around the memory-prediction model of Hawkins and Blakeslee (2004) as outlined in the book 'On Intelligence. The core focus of this model is a simplified version of the neocortex, with the same hierarchical processing structure is found across the neocortex and so all regions have the capacity to perform the same activities [Moore 2007, Garalevicius 2007]. For Hawkins and Blakeslee (2004) the neocortex works by storing sequences that are used to determine what will happen in the person's environment [Moore 2007] and make predictions through various associations. The lowest level of the neocortex receives inputs from the senses and represents these using temporal and spatial methods and the neocortex learns series of patterns by storing a representation of them using a hierarchical structure [Garalevicius 2007]. Attention is controlled in the model by the novelty associated with a stimulus with an unusual occurrence in the hierarchy eventually becoming part of the consciousness [Moore 2007]. The areas of the hierarchical structure are linked through feedforward and feedback connections to make predictions by contrasting the information from the feedforward and feedback connections [Moore 2007, Garalevicius 2007]. Each area in the neocortex is made up of a set of subareas and is only connected to another area through areas higher up in the hierarchical arrangement.

Garalevicius et al. (2007) developed a Bayesian model for visual pattern recognition based on the memory-prediction model using the approach of George and Hawkins (2005). The hierarchical structure of the model has various subareas, with the lower layer (level 0) partitioning the image into sections and the upper layer (level 2) producing a prediction of what the visual image is. The subareas in level 0 memorise each of the input patterns and the likelihood of them occurring above a specific threshold as the output to a subregion of level 1. Each pattern is represented by an index value called its 'name'. From these 'names' level 1 learns the most likely pattern combinations from the subareas linked to it in level 0 and passed its 'name' back to the active patterns in its connected subareas in level 0. A specific image is explained by a group of patterns that have the highest probability value through Pearl's Bayesian belief revision algorithm. This enables the subareas in lower level to produce a conditional probability distribution matrix to represent a probability of the pattern occurring in this area given the pattern in the subarea in the level above.

In Garalevicius et al. (2007) model, learned patterns for level 0 are preset and made up of 139 primitives of typical observed stimuli such as lines and corners, with a class tag such as 'top left corner'. The section of the image to the subarea in level 0 is compared to the stored patterns and the stored pattern with the lowest Hamming distance is chosen. The class tag is then used as the 'name' of the stimulus. The upper region of the memory model has a fixed set of possible predictions to select from as the final prediction of the network. The model is trained on 91 categories of black and white drawings, with each category containing 2 drawings. An unrelated person redraw the training data to act as the test images. Garalevicius et al. (2007)'s model was tested using a variable number of saccades, with a saccade being represented by moving each image one pixel diagonally, and found to have good autoassociator capabilities by dealing with missing components of the pictures. To model the capacity to forget patterns the frequency of patterns in the training set are determined and those patterns that occurred rarely are removed. Contextual information related to the most commonly occurring patterns is passed back to create conditional probability matrices. The model developed uses an interference approach which offers the advantage that additional categories could be added and further training could occur after the original training process. The memory-prediction model provides an overall model of the human neocortex that can form the basis of many models and predictive systems, which makes it suitable for use in the memory model to be developed in workpackage 3 as part of the ACORNS project.

3 Memory models of attention mechanisms

It is likely that the memory model developed for ACORNS project will make use of the cognitive memory function of attending to specific stimuli at the expense of others. In our every day life we are faced with a constant input of stimuli from multimodal sources. It is not possible for the brain to process all of these inputs at the same level and so the brain uses approaches that bias attention to specific events [Kayser et al. 2005]. Hence, attention is the procedure of focusing on one component of our surroundings at the expense of others. An example of attention is the cocktail party affect where we concentrate on listening to one person and hence ignore other noise and conversations in the room [Arons 1992]. However, attention is an extremely complex cognitive function which has caused many disagreements in the neuroscience community when determining an appropriate definition and what it actually incorporates [Pugh et al. 1996]. For Pugh et al. (1996) attention should include the capacity to switch focus from one element to another, must be maintained over a period of time and be limited in the number of elements that can be focused on at any time.

3.1 Biological basis of attention mechanism

Various neuroscience studies have considered aspects of attention such as sustained attention, selective attention and decision and action control [Pugh et al. 1996]. For sustained attention the brains right hemisphere has a more active role than the left hemisphere. Pardo et al. (1991)

showed that there is activation found in the superior parietal area and the prefrontal areas of the right hemisphere when subjects looked for small changes in stimulus [Pugh at al. 1996]. Turning to selective attention the superior parietal lobule is identified to be related to removing focus from one element to the next, the superior colliculus the movement to the new focus and pulvinar for filtering out stimulus that are not of interest [Posner and Presti, 1987, Posner and Petersen,1990, Pugh et al. 1996]. Attention associated with decision and action control appears to be performed by the anterior brain area [Pugh et al. 1996]. Corbetta et al. (1991) when requiring subjects to study a display for characteristics there was greatest activation in the cingulate when selective attention is required. When considering selective attention in the auditory system Pugh et al. (1996) found that there is an increase in the activation in the parietal lobule. Although the activation is greatest in the inferior partial lobe increased activation was also found to an extent in superior and parietal lobule and precuneous areas. According to Zatorre et al. (1999) selective auditory attention comes from both spatial and frequency cues. Zatorre et al. (1999) used an auditory attention task to test which regions are activated when the participants processed tones of different frequencies in the left, right or both ears. Regions associated with auditory attention were found predominately in the right hemisphere in the right parietal, frontal, and temporal cortex.

According to Nieuwenhuis and Yeung (2005) there are two current theories related to attention. The first is that the brain suppresses those cognitive activities not associated with the task that is being focused on and the second is that the prefrontal cortex has the capacity to inhibit stimuli that is not currently relevant. However, it is still disputed whether inhibition is the main approach by the prefrontal cortex to achieve attention. In response to these theories Egner and Hirsch (2005) considered whether the prefrontal cortex uses amplification or inhibition to aid attention through a variation of the stroop task. In the stroop test subjects are required to state the colour of the printed colour word. Performance is usually poorer when the colour and the words are different, for instance when the word reads red but is blue in colour [Nieuwenhuis and Yeung 2005]. The variation by Egner and Hirsch (2005) replaces print colour with a faces, with the incorrect name of a politician or actor being written on a photograph of an actor or politician face. Half of the subjects are advised to concentrate on one feature (the photograph) and the other half on the other feature (written name). The subjects are then required to state if the attribute they are required to concentrate on is of a politician or actor. In some of the cases the text and photograph match in terms that they are both of either an actor or politician in others they differ. According to Nieuwenhuis and Yeung 2005 the results found in this experiment supporting the view the fusiform face area is involved in memory activity amplification of the relevant information. When concentrating on the photograph there is greater activation in the fusiform face area but when concentrating on the words there is no reduction in activation in the fusiform face area despite the photograph being irrelevant. However, as noted by Nieuwenhuis and Yeung 2005, there is no indication by Egner and Hirsch (2005) how the prefrontal cortex understands what information is relevant and what is not.

3.2 Computational memory models of attention

For Kayser et al. (2005) attention can be seen as biasing toward a specific event, with elements in the environment weighted based on their current importance. Which stimuli are considered for full examination depends on an involuntary and stimulus-based approach and a cognitive element that involves voluntary control. The stimulus-based approach provides a weighted depiction of the environment dependent on those stimuli that are related to an event. To comprehend how salient stimuli are chosen the saliency map approach was used by Kayser et al. (2005) to create a model of attention memory in the auditory system. Saliency maps in this approach are seen as models that retrieved characteristics in a parallel hierarchical manner. The auditory system segmented sounds from a complicated scene based on characteristic such as frequency, pitch and amplitude to represent temporal, spatial and spectral information so as to offer the opportunity to identify interesting sounds against noisy backgrounds [Winter and Schreiner 2005]. In Kayser et al. (2005) approach the saliency map retrieves features such as intensity, frequency contrast and temporal contrast using filters on different scales. For every feature, maps are compared using a central-surrounding approach and normalised to identify peaks. The central-surround maps are combined by collapsing the scales to produce the saliency map for the specific feature. These saliency feature maps are combined to produce the finally map (see Figure 3.2.1). This model is used in the visual system and as such indicates that stimulus-based attention memory is similar across different modalities.

A further model of attention memory based on saliency maps is that of Choi et al. 2004, who considered the visual system. This model offers a trainable attention approach that allows the inhibition of salient areas that are not seen as of interest. The model is implemented using a bottom up saliency map based on four features: intensity, edge detection, colour and symmetrical information. Top-down an adaptive resonance theory network is trained based on a human expert to recognise/memorise areas of the map that are not of interest and ignore them in future saliency maps. When developing the saliency map, maps are produced for the four features using central surround difference and normalisation. Once these maps are produced they are combined using independent component analysis filters. Although salient areas are identified, such areas might not be of interest to humans. In response to this a selective attention memory approach is devised which like humans ignored areas despite them being described as salient based on primitive features such as colour, edge etc. This is achieved using the adaptive resonance theory by interaction with a human expert who provides information on the salient regions that are not of interest, which are inhibited by the model [Figure 3.2.2].

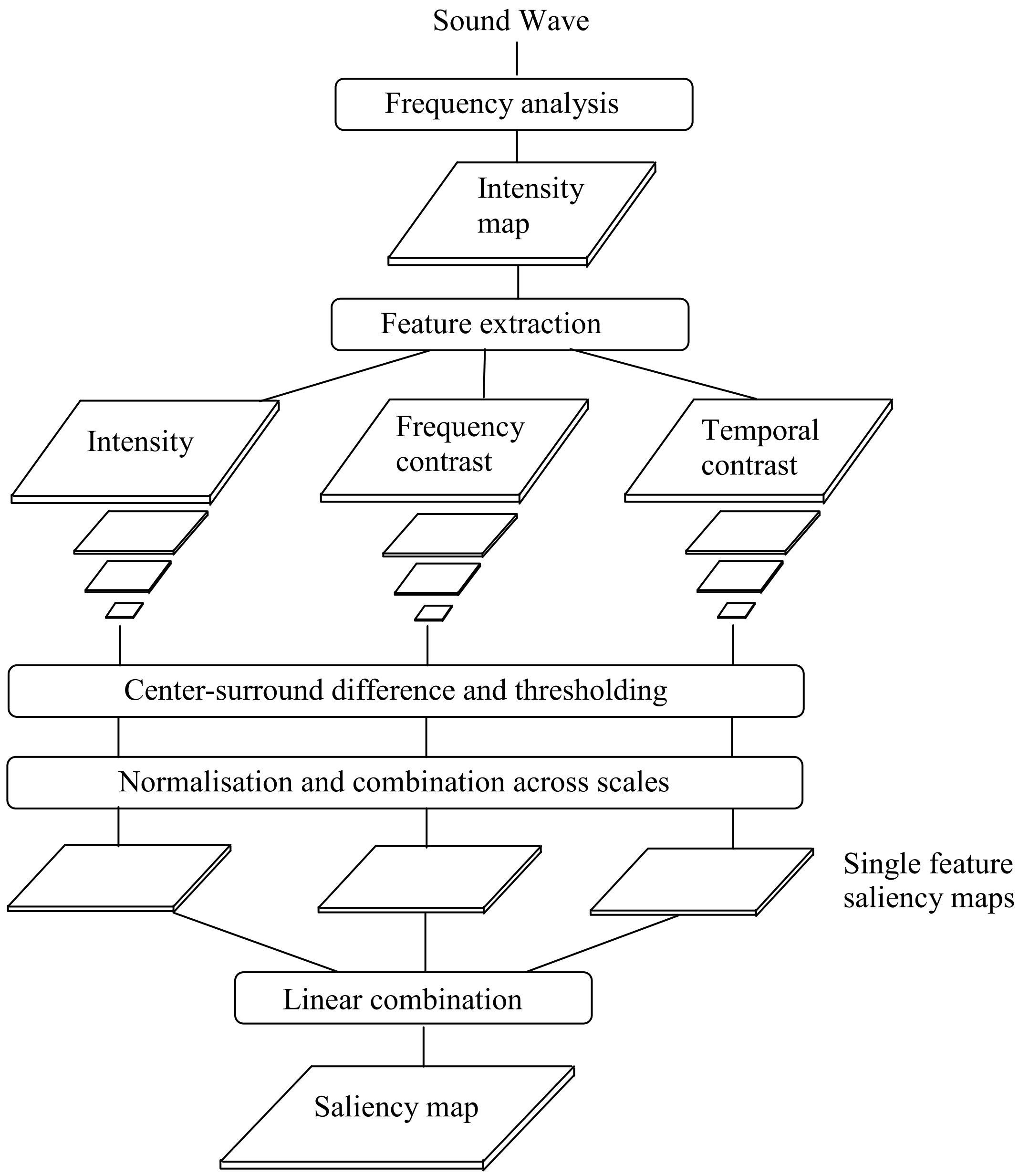


Figure 3.2.1 Model of the auditory attention mechanism based on saliency maps (After Kayser et al. (2005)).

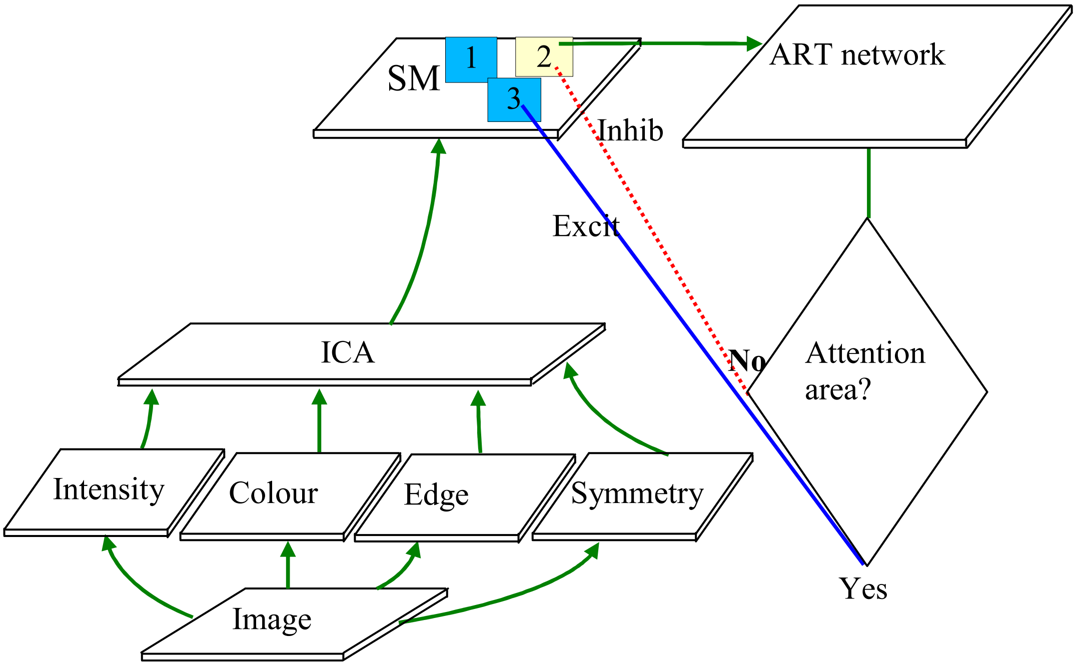


Figure 3.2.2 Visual attention system (After Choi et al. (2004))

A computational model of attentional memory was also developed by Iwasaki et al. (1999). It determined features of interest and areas of attention using a hierarchical neural network [Figure 3.2.3]. The network uses a three layer model to establish the attention areas, the associator layer, the centre layer and the symbolic layer. The association layer is used to identify relevant features of the stimulus and the center layer acts as the feature extraction layer. The units in each layer are linked to those in the other layers with the weights being updated using the Hebbian learning rule. The weights between the input and centre layer give a basic depiction of the stimulus and the links between the symbolic and input layers depiction the attention regions. This approach uses an interaction between bottom-up and top-down methods to achieve the learning of attention. In the bottom-up approach the stimulus is passed to the associator layer and then passed up to the centre and symbolic layers. In the top-down method units in the associator layer are selected and the associator layer is driven using these units. The bottom-up, top-down process is repeated until the outcomes are the same.

The learning procedure for the hierarchical neural network involves the following: (a) an input pattern is introduced to the associator layer; (b) the most excited units in the centre c_i and the symbolic layer s_j are activated by the bottom-up process. (c) A section of neurons in the input layer are selected for the top down process; (d) A section of the associator network is selected to drive this layer ; (e) The most excited units $c_{i'}$, $s_{j'}$ are activated using the bottom-up process; (f) if the most excited neurons for the centre and symbolic layer are the same for the full associator layer as for the smaller area of associator layer ($c_i = c_{i'}$ and $s_j = s_{j'}$) the links between the units are updated using the Hebbian learning rule, (g) or else, another portion of the input is chosen and the process repeat until they match; (h) steps (a) to (g) are performed for all inputs; (i) steps (a) to (h) are performed until the stop requirement is fulfilled.

The weights between the units in the layers are updated using Hebbian learning algorithm. The learning approach presented by Iwasaki et al. (1999) is as follows:

$$w_{AC_{ij}}^{new} = w_{AC_{ij}}^{old} + \Delta w_{AC_{ij}}$$

$$w_{AS_{kj}}^{new\ new} = w_{AS_{kj}}^{old} + \Delta w_{AS_{kj}}$$

$w_{AC_{ij}}$ represents the connect weight between the associator layer and the centre layer. $w_{AS_{kj}}$ represents the connection weight between the associator layer and the symbolic layer. $\Delta w_{AC_{ij}}$ and $\Delta w_{AS_{kj}}$ are the update values for the selected neurons in stage (d) above.

$$w'_{AC_{ij}} = \frac{W_{AC_{ij}}}{\sum \left| W_{AC_{il}} \right|}$$

$$w'_{AS_{kjj}} = \frac{W_{AS_{kjj}}}{\sum \left| W_{AS_{kl}} \right|}$$

$$w_{CS_{ik}}^{new} = w_{CS_{ik}}^{old} + \Delta w_{CS_{ik}} - \beta$$

$w_{CS_{ik}}$ represents the connect weight between the centre layer and the symbolic layer. The weights are updated using the Hebbian rule, with a constant β fixing the degree that the model forgets. $\Delta w_{CS_{ik}}$ is updated based on the most excited neuron.

At first the weight values for the links between the centre layer and the associator layer are random. The updating of these uses Hebbian learning based on a winner-take-all approach. The connection weights between the centre layer and the symbolic layer are reduced by a forgetting constant if they are activated. To prevent units that have no activation, if any unit in the centre layer have no activation it is removed and replaced with a unit whose connection weights are the same as the unit that is most often active. In the neural network the weights in the centre and associator layers and the weights between the symbolic and associator layer are updated based on common details on the characteristics of the input based on Hebbian learning. As a result this may produce the situation where the same characteristic is represented by both weights. If this is the cases the appropriate weights between the symbolic and associator layer are reduced by a fixed amount.

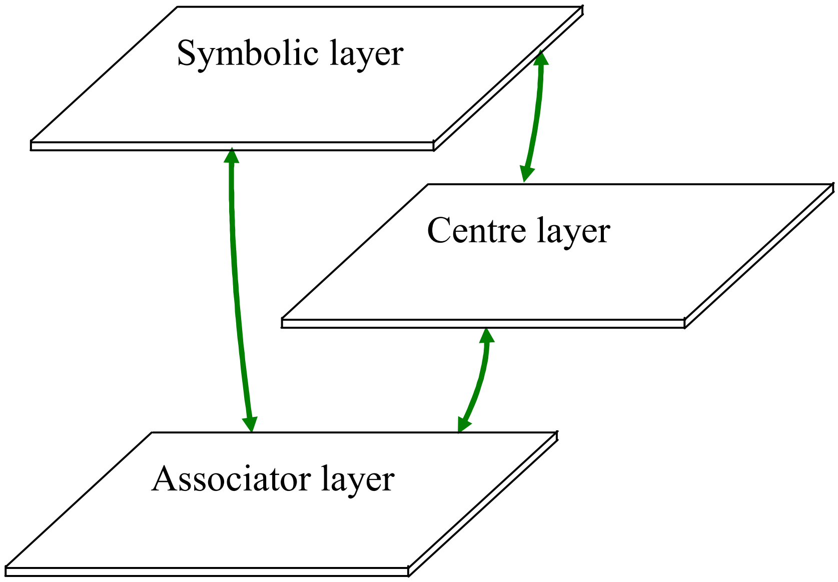


Figure 3.2.3 The hierarchical neural network for attention (After Iwasaki et al. (1999))

The final attention computational memory model considered here relates to auditory attention was developed by Wrigley and Brown (2002, 2004). This model of attentional memory incorporates three steps: sampling of the audio signal using two channels; retrieval of periodic information using a correlogram that created a 'binaural' F0 estimate and combination and segregation that used a neural oscillator network based on locally excitatory and globally inhibitory oscillation. Each oscillator passes activity to the attentional leaky integrator which is central to auditory segmentation and produces the attentional motivation stream. The weights between the oscillator network and the attentional leaky integrator are altered by endogenous procedures.

The individual oscillators are represented in the model by x and y:

$$\dot{x} = 3x - x^3 + 2 - y + I_o$$

$$\dot{y} = \varepsilon[\gamma(1 + \tanh \frac{x}{\beta}) - y]$$

ε, γ and β are parameters. Oscillations are only active when $I_o > 0$. The input I_o to oscillator i is made up of external input I_r, network activations and global inhibition.

$$I_o = I_r - W_z S(z, \theta_z) + \sum W_{ik} s(x_k, \theta_x)$$

W_{ik} are links between the oscillators i and k and x_k is the activity of the oscillator k. θ_x acts as a level above which an oscillator could influence other oscillators and w_z are the weights for the global inhibitor z. θ_z is the level above which the global inhibitor could influence the oscillators. S acts as a squashing function.

The model of attentional memory by Wrigley and Brown (2002, 2004) uses a set of gammatone filters to recreate the frequency selectivity associated with the basilar membrane, with each filter acting as the different points along the membrane. The correlogram in the system is able to determine formant and harmonic regions that are used to produce segments. A segment is produced by determining the cross correlation between channels of the correlogram and from this cross correlation value producing a similarity level to determine if channels are similar enough to be grouped into a segment. The segment estimation procedure therefore happens in two sections. First, there is the determination of periodic segments which are the channels that have a peak in the energy weighted inverse variance which is greater than a specific tonal level. Second, there is the determination of the noise segments which are those remaining peaks that are greater than a certain noise level. In this model segments are grouped based on common harmonicity by using additional excitatory links between the oscillators in the segments.

An oscillatory is linked to the attentional leaky integrator using an excitatory connection based on the level of endogenous attention. The response of the attentional leaky integrator gives the frequency component of the attentional stream at a particular time step. An input to the attentional leaky integrator is:

$$ali = H\left(\sum S(x_k, \theta_x) T_k - \theta_{ALI}\right) - ali$$

θ_{ALI} acted as threshold above which activity from network influences the ALI. T_k performed as an attentional weighting associated with the endogenous interest at frequency k:

$$T_k = 1 - (1 - A_k)L$$

According to Wrigley and Brown (2004) A_k acts as the endogenous interest at frequency k and L is the leaky integrator.

$$\dot{L} = a(b[R - L]^+ - [1 - H(R - L)]fL)$$

Attentional interest is represented using a Gaussian

$$A_k = \max_{A_k} e^{-\frac{k-p}{2\sigma^2}}$$

A_k acts as the normalised attentional interest for frequency channel k and $\max_{Ak}$ is the largest value that A_k can achieve. p is the channel that as the peak of attentional interest, and s is the width of the peak.

The model of Wrigley and Brown (2002, 2004) uses an attentional leaky integrator and a depiction of attentional allocation between frequencies, which matches the preference of the person listening and is represented using Gaussian distributions. This attentional memory model is seen to produce a number of auditory grouping occurrences where attention is seen to be critical.

4 Sensor/motor memory model for grounding of language

In this section of the report, we consider learning/memory models for the grounding of language using sensor/motor representations in terms of biological basis and those computational models that have been developed. Such sensor/motor memory models are to be incorporated in the ACORNS memory model as it is fundamental that language acquisition has the capacity to ground meaning in the symbols or vocal sounds that constitute language [Vogt 2006]. Grounding is the link between symbols or vocal sounds in the form of words and the physical world and hence associates high level language with low level sensor information to refer to actions and objects which is fundamental for communication [Roy 2004, Sun 2000]. Harnad (1990) and (2003) described the concept of the symbol grounding problem in that abstract symbols or vocal sounds must be grounded or be associated to something in the real world to interpret their meaning. Hence, to actually attribute meaning to a word there must be interactions with the world to provide relevance to the symbolic representation and so not describe words simply in terms of other words [Roy 2003].

4.1 Mirror neuron system

The first biological basis of sensor/motor memory consider here is the mirror neuron system. Rizzolatti and Arbib (1998) and Umilta et al. (2001) found that neurons located in the rostral region of a primate's inferior area, the F5 area [see Figure 3.1.1] are activated by the movement of the hand, mouth or both. These neurons fire as a result of the action not the movements that are the components of this action. The recognition of motor actions came from the presence of a goal and so the motor system does not solely control movement [Gallese and Goldman 1998, Rizzolatti 2002]. Hence, what turns a set of movements into an action is the goal and holding the belief that performing the movements would achieve a specific goal [Arbib 2005]. The F5 neurons are organised into diverse categories based on the action that cause them to fire, which are 'grasping', 'holding' and 'tearing' [Gallese and Goldman 1998, Rizzolatti and Arbib 1998].

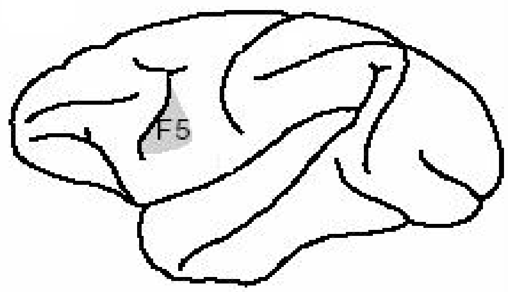


Figure 4.1.1 A representation of the cerebral cortex showing the rostral region of a primates inferior area, the F5 area.

Certain grasping-related neurons fire when grasping an object whether the grasping is performed by the hand, mouth or both [Gallese et al. 2004]. This supported both the view that these neurons do not represent the motor action but the actual goal of performing the grasping task. Within area F5 there are two types of neuron: the first, known as canonical neurons, only respond to the performing of the action and the second mirror neurons that respond not only when performing an action but also when seeing or hearing the action performed and so these primates have developed a memory associated with the behaviour [Kohler et al. 2002, Rizzolatti and Arbib 1998, Rizzolatti et al. 2001]. The mirror neuron system indicates that the motor cortex is not only involved in the production of actions but in the action understanding and memorisation from visual and auditory information [Rizzolatti et al. 2002, Rizzolatti and Luppino 2001, Rizzolatti et al. 1998] and so the observer has the same internal memory representation of action as the actor [Umilta et al. 2001].

These mirror neurons are typically found in area F5c and do not fire in response to the presence of the object or mimicking of the action. Mirror neurons require the action to interact with the actual object. They respond not only to the aim of the action but also how the action is carried out [Umilta et al. 2001]. However, as shown by Umiltà et al. (2001) an understanding that the object is there without being visible causes the activation of the mirror neurons if the hand reaches for the object in the appropriate manner. This is achieved when primates are first shown the action being performed completely visible and then with the hand-object interaction hidden. As the performance and recognition of an action causes activation in the premotor areas which are responsible for the hand movements when simply observing the action there is a set of mechanisms that suppress the movements to perform the action.

Given the nature of the mirror neuron system one possible scenario is to achieve the grounding of perceptual information in actions through imitation learning. According to Schaal et al. (2003) imitation learning is able to speed up the learning process. As the in the ACORNS project there is interaction between an intelligent agent and carer some form of imitation learning might have a role in the language acquisition process. Imitation learning allowed the observer to gain skills by creating an abstract memory representation of the teacher's behaviour, understanding the aims of the teacher and creating the solution [Dillmann 2003]. Imitation required the ability to take the seen action and produce the appropriate motor primitives to recreate this [Buccino et al. 2004]. The mirror neuron system is held to have a major role in immediate imitation if the action that occurs is in the observer's repertoire/memory [Buccino et al. 2004]. The role of mirror neurons is to depict actions so they are understood or can be imitated, by gaining the reason for the action [Rizzolatti and Arbib 1998, Sauser and Billard 2005].

A possible explanation for the ability to imitate is the internal memory vocabulary of actions that are recognised by the mirror neurons [Rizzolatti and Luppino 2001]. This ability to understand others actions, beliefs, goals and expectation aids the inclusiveness of the group. This allows the observer to predict the future actions and so determine if they are helpful, unhelpful, threatening and to act accordingly [Gallese and Goldman 1998, Gallese 2005]. It is argued by Demiris (2002) and Demiris and Hayes (2002) that through the mirror neuron system when a primate or human watches an action they are to imitate they put themselves in the place of the demonstrator. Understanding the actions of the demonstrator comes from creating alternatives and choosing the most appropriate one. The ability to predict the action rather than waiting until it is complete offers the opportunity to react to the action before it has ended. A requirement for imitation is to connect the sensory system with the motor system so that the multimodal inputs are linked to the appropriate actions.

As a response to the theories related to mirror neurons there has been a considerable amount of research into the mirror neuron system and in particular for imitation. For instance, Buccino et al. (2004) study using event-fMRI the cerebral cortex activations produced when individuals who could not play the guitar watch another person's finger positions for specific cords and are requested to memorised and recreate these finger positions. The outcome of this experiment shows that the human mirror neuron system is central in memorising behaviour using imitation with activation in the rostral part of the inferior partietal lobule and the ventral premotor cortex as well as the pars opercularis of the inferior frontal gyrus.

4.1.1 Language and the mirror neuron system

We consider the role of the mirror neuron system in the evolution of language which specifically relates to the ACORNS project and the role of the mirror neuron system in the emergence of language. Turning to the human mirror neuron system, it is observed that mirror neurons in humans are also excited by both the performance and observation of an action [Gallese and Goldman 1998]. The F5 area in primates corresponds to various cortical areas in humans including the left superior temporal sulcus, the left inferior parietal lobule and the anterior region of Broca's area. The association of mirror neurons with Broca's area in humans and F5 in primates provides an indication that mirror neurons might have evolved in humans into the language system [Rizzolatti and Arbib 1998]. The role of the mirror neuron system in language can be seen from the findings of Pulvermüller [Hauk and Pulvermüller 2004, Pulvermüller 2003] in that the processing and memorised representation of words includes the activation of some of the same regions as those that are found to perform the action. The ability in the first instance to recognise an action is required for the development of a communication system between members of a group and finally for an elaborate language system [Kohler et al. 2002, Rizzolatti and Arbib 1998].

In order to study whether action creation and language production are linked, Hamzei et al. (2003) performed an fMRI study that required the performance of action recognition, language production and grasping movement. This study found that action recognition regions of the cerebral cortex are found in the left inferior frontal gyrus and between the inferior frontal gyrus and precentral gyrus, the ventral occipito-temporal junction, the superior and inferior partietal cortex, and the intrapartietal sulcus in the left hemisphere. There is an overlap of activation produced by action recognition, language production and action production in the partial cortex, the left frontal gyrus, and the inferior frontal gyrus and precentral gyrus border. Arbib (2005) notes various studies that also support the idea that observation and preparing to perform grasping both included cerebral cortex regions that are associated with speech production. Gerlach et al. (2002) states that the left ventral motor cortex is activated for words related to tools, food and clothing, animals and non-manipulative man-made objects, which points to the fact that certain lexical categories develop from action-based knowledge.

Arbib (2005) discussed the role played by the mirror neuron system in language evolution and so the memorised grounding of language in actions. He examined the neural and functional basis of language and the evolution from the action recognition and production performed by primates to the full language in humans. The article points to the belief that the mirror neuron system is integral in the development of the language ready cortex and the multimodal nature of language. The F5 mirror neurons in monkeys are connected to areas of the parietal and temporal cortex and through evolution the F5 region becomes Broca's area and these other areas become Wernicke's area and other language regions. Arbib (2005) states that language development involved 7 stages: (i) grasping; (ii) a mirror neuron system for grasping; (iii) a simple imitation system for object grasping; (iv) a complex imitation system that allows the grasping action to be recognised and then repeated; (v) a gesture based language system; (vi) protospeech; and (vii) language that moves from action object frames to a semantic syntax based approach.

Arbib (2005) points out the strong inputs to the F5 area neurons from the secondary somatosensory area and parietal area PF. The canonical neurons of F5 area are also the selective targets of the anterior intraparietal cortex. Many visual response neurons in PF are found to discharge when observing and performing an action and hence are known as area PF mirror neurons. This connection arrangement enables an observation/execution system with mirror neurons in area F5 and area PF firing when performing an action and observing it. The extension of the mirror neuron system in humans allows complex imitation that aids production of compound actions and the development of language. The ability of humans to engage in pantomime behaviour enables a vocabulary of gestures to occur and so a gesture communication system. By use of a gesture based system it allows the development of an open communication system from fixed vocalisations by primates. According to Arbib (2005) this gesture communication system moves from imitation to perform an action to imitation for communication and is fundamental for the development of a spoken vocabulary approach known as protospeech. Protospeech becomes associated with the gesture system and once this occurs the gestures are no longer required and so they reduce in importance or disappear.

Protolanguage is seen to consist of a set of unitary utterances that can provide a great deal of information and a precursor to a full language system. A single protolanguage utterance according to Arbib (2005) could mean “the animal has been killed with a spear take the meat to the front of the cave”. Such an approach however can suffer from a situation that there is a need for a large number of different words to represent all the information that is needed. For instance, if there is only a single word for each of ‘ripe apple’, ‘sour apple’, ‘sweet apple’, ‘sweet pear’, ‘ripe pear’ and ‘sour pear’ this requires a lot more word storage than if there is separate words. Hence, a likely development in the protolanguage is the use of multiple words. Through the development in language capabilities there is finally a move from protospeech to language which is achieved through the development of named actions, a complex syntactic and semantic structure, the identification of the hierarchical structure of language and the employment of verb-argument structures.

Hence, evolution has enabled humans to move from the basic mirror neuron system that memorises and recognised actions to a complex language system that allows cultural development. According to Arbib (2005) this evolution is not achieved by the replacement of one capability with another one, but with the adding of capabilities within an existing system. Such a language system offers a common understanding by both the speaker and the listener [Arbib 2005].

4.1.2 Computational models of mirror neuron system for grounding

Considerable interest has been expressed in the use of the mirror neuron system principle as the basis of grounding [Belpaeme et al. 2003]. Based on the mirror system hypothesis, Billard and Matarić (2001) developed a robot capable of imitation. Their approach uses a hierarchy of neural networks and provides an abstract and high level depiction of the neurological structure that is the basis of the visuo-motor pathways to examine the ability to reproduce human arm movements. This memory based model consists of three parts for visual recognition, motor control and learning and uses seven modules. A module based on the temporal cortex processes visual information to identify the direction and orientation of the teacher's arms with reference to a point on the teacher's body. The motor control is based on a hierarchical model with a spinal cord module at the lower level. Learning/memory of movement occurs in the premotor cortex and cerebellum modules and learning creates links between the primary motor cortex, premotor cortex and the cerebellum and within the premotor cortex and the cerebellum. These modules use a dynamic recurrent associative memory architecture which is a fully connected recurrent network that enables time series and spatio-temporal data to be learned using short-term memory. The model when tested on series of arm movements is found to reproduce all motions despite the noisy environment.

The dynamic recurrent associative memory architecture approach is also used by Billard (1999) in two experiments for grounding a protolanguage by using the mirror neuron principle. By using the dynamic recurrent associative memory architecture recurrent memory approach the student robot is able to learn actions and the labels associated with them. In the first experiment a wheeled student robot learns a protolanguage by following the teacher around and associating/memorising radio signals as a form of language from the teacher robot with different sensory inputs that provide the location of landmarks. The input modalities in this case are the robot’s own location and a form of language. Furthermore, a small doll robot uses the dynamic recurrent associative memory architecture to imitate the arm and head movements of the teacher. The doll robot is trained to perform a series of leg and head movements by the human teacher based on the readings from infra-red sensors giving the location of the teacher’s head or arms and labels typed in by the teacher.

Amit and Matarić (2002) also perform memory grounded learning of movement sequences based on the mirror neuron system principle using a hierarchical structure. A teacher who is either a robot or a human performs an action and this is learned through imitation by a robot. The approach uses a hierarchical framework made up of three layers: base primitives, movement specialisers and sequence learners. Each base primitive is the visio-motor primitive that encodes the motor programs to perform a class of movements. The base primitives are felt to be innate and so are manually encoded. The next level consists of the movement specialisers that specialise on observed movements. Movement specialisers are learned/memorised to represent movements. Each base primitive represents a generic class of movements such as reaching, however each movement specialiser learns a particular movement such as reaching for a cup on the table. The sequence learning is the highest level of the hierarchy and represents complex movements by combining a set of movement specialisers. In the model the sequence learning is produced by using Hidden Markov Models. The aim of the experiment for the robot is to learn 8 aerobic behaviours based on movements of the arms. The system uses a base primitive for each of the two shoulders, 30 specialisers for the movement specialiser layer and 10 Hidden Markov Model based sequences learners. The experiments involve presenting the aerobic sequence with the movement specialisers learning what is happening online as the demonstration occurs.

Another system for grounding based on the mirror neuron system is that of Tani et al. (2004). The approach uses a recurrent neural network with parameter biases (RNNPB) for both recognition and production of behaviours. In this approach, sections of spatio-temporal data of sensory-motor flow are depicted by using vectors of small dimensions. The nonlinear dynamical system is produced using a Jordan-type recurrent network that had parametric biases (PB) incorporated in the input layer function. In order to reproduce the characteristics of the mirror neuron system, the RNNPB creates the appropriate dynamic pattern from fixed PB and performes recognition by producing the PB from a target pattern. There is learning/memorising of movement patterns using the forward model by producing the PB vectors and a syntactic weight matrix. Following learning it is possible to produce sensory-motor series by using the forward dynamics of the RNNPB with the parameter biases fixed. When the network produces a behaviour it operates in a closed loop where the prediction of the next action is fed back as an input.

In the experiments, a Sony humanoid QRIO SDR-4XII robot attempts to produce synchronised hand movements with a human after first learning/memorising the hand movements. When a hand movement is identified by the robot it alters the PB to reduce the sensory prediction errors. The RNNPB has 12 input units and 12 output units to learn the forward dynamics of movement patterns. The patterns contain information on the position of the hand and the angle of the joints of the robot arms and so allowed the robot to ground the action in the environment. There are also 4 parametric units, 40 hidden units and 30 context units. In the first experiment, the human user trains the robot with three different behaviours over a period of 20 seconds. It is found that the robot is able to synchronize well with the human and when at one point it did lose synchronization it regained it. In this experiment the robot is shown to produce emergent behaviour when the human produces hand actions that the robot had not seen before.

The RNNPB memory model is also used for grounding of language in actions by Tani et al. (2004) based on the mirror neuron system. As it can be seen from Figure 4.1.2.1 this is achieved by using a language system that learns word sequences using a network with 10 input units, 10 output units, 6 PB units, 50 hidden units and 4 context units and a sensory-motor network that learned behaviours that incorporates 26 input units, 26 output units, 6 PB units, 70 hidden units and 4 context units. The model uses the word sequence and sensory-motor sequences as input at time t and tried to predict these values at $t+1$ and the PB values. The models use the actual values at time $t+1$ as feedback to backpropagate through the system. During learning the PB_{lang} and PB_{behav} are updated with the aim to reduce the difference between these PB values. The word sequences are grounded to the appropriate sensory-motor sequences using these PB values. The robot learns to perform the actions 'hit', 'push' and 'point'. The word sequences are typically made up of two words: a verb and a noun. The verb was either 'hit', 'push' or 'point' and the nouns are 'red', 'green', 'blue', 'left', 'centre' and 'right'. In this case 'red', 'green' and 'blue' referred to blocks that are to the 'left', 'centre' and 'right' respectively. It is possible to generate goal-directed actions by identifying two word phrases.

Motivated by the mirror neuron model, Wermter et al. (2005) develop a grounding in language memory based approach which allows a student robot to learn from the teacher robot who performs three behaviours 'go', 'pick' and 'lift' based on multimodal inputs. To allow the student robot to learn/memorise these behaviours, the teacher robot performs 'go', 'pick' and 'lift' actions one after another in a loop in an environment.

The student robot observes the teacher robot performing the behaviours and was trained by receiving multimodal inputs. These multimodal inputs are: (i) high-level visual inputs which are the x- and y-coordinates and the rotation angle φ of the teacher robot relative to the nearest wall; (ii) the motor directions of the robot ('forward', 'backward', 'turn left' and 'turn right'); and (iii) a language instruction stating the behaviour the teacher is performing ('go', 'pick' or 'lift').

The docking procedure performed by the teacher that is involved in the 'pick' behaviour is based on a reinforcement learning actor-critic approach of Weber et al. (2003) and Weber et al. (2004) using the research of Foster et al. (2000) [See Section below on actor-critic model]. The final action verb 'lift' involves moving backward to leave the table and then turning around to face toward the middle of the arena. Coordinates x and φ determines how far to move backward and in which direction to turn around. The robot moves backwards until it reaches a specific point on the x-axis and then turned to a random angle to face the back wall. If the angle that the robot turned to is negative the robot turns to the left, if it was positive it turned to the right.

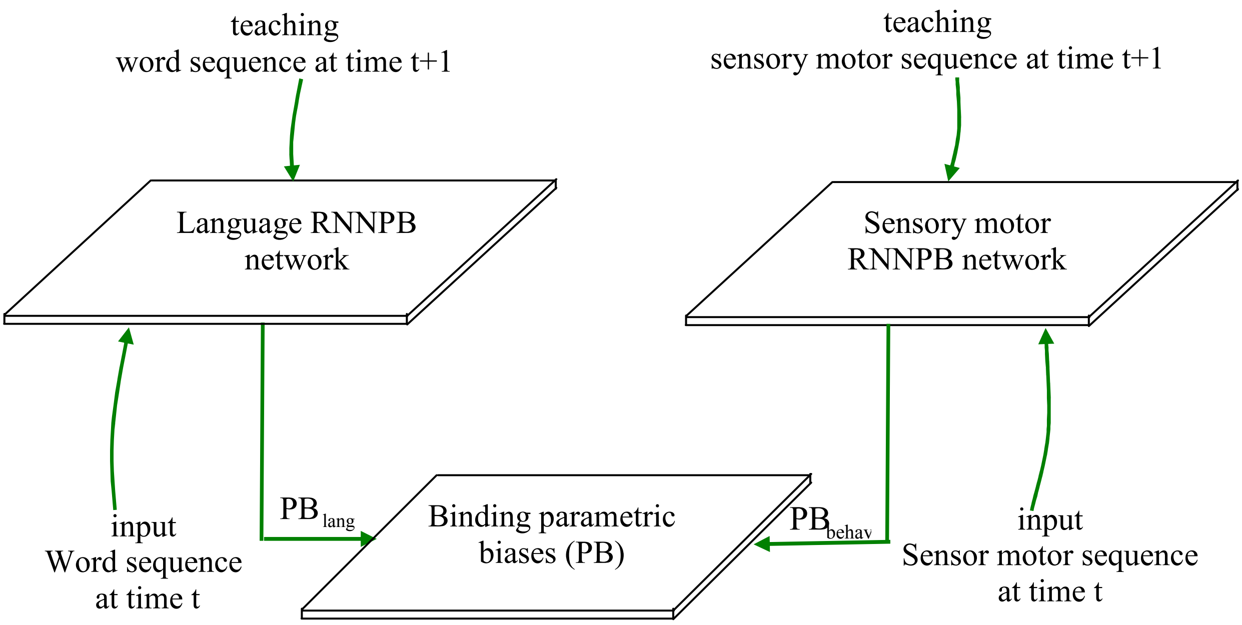


Figure 4.1.2.1 The RNNPB based model for grounding language in actions (After Tani et al. 2004).

Motivated by the mirror neuron model, Wermter et al. (2005) develop a grounding in language memory based approach which allows a student robot to learn from the teacher robot who performs three behaviours 'go', 'pick' and 'lift' based on multimodal inputs. To allow the student robot to learn/memorise these behaviours, the teacher robot performs 'go', 'pick' and 'lift' actions one after another in a loop in an environment.

The student robot observes the teacher robot performing the behaviours and was trained by receiving multimodal inputs. These multimodal inputs are: (i) high-level visual inputs which are the x- and y-coordinates and the rotation angle φ of the teacher robot relative to the nearest wall;

(ii) the motor directions of the robot (‘forward’, ‘backward’, ‘turn left’ and ‘turn right’); and (iii) a language instruction stating the behaviour the teacher is performing (‘go’, ‘pick’ or ‘lift’).

The docking procedure performed by the teacher that is involved in the ‘pick’ behaviour is based on a reinforcement learning actor-critic approach of Weber et al. (2003) and Weber et al. (2004) using the research of Foster et al. (2000) [See Section below on actor-critic model]. The final action verb ‘lift’ involves moving backward to leave the table and then turning around to face toward the middle of the arena. Coordinates x and φ determines how far to move backward and in which direction to turn around. The robot moves backwards until it reaches a specific point on the x-axis and then turned to a random angle to face the back wall. If the angle that the robot turned to is negative the robot turns to the left, if it was positive it turned to the right.

When receiving the multimodal inputs relates to the teacher's actions the student robot is required to learn/memorise these behaviours so that it could perform them from a language instruction or recognise them. When learning, performing and recognising these behaviours ‘forward’ and ‘backward’ movement is at a constant speed of 1 unit for each time step and the decision to ‘turn left’ or ‘right’ is 15^o each time step. A neural memory model hierarchical architecture [Figure 4.1.2.2] uses an associator network based on the Helmholtz machine learning approach [Dayan 2000, Dayan and Hinton 1996, Dayan et al. 1995, Hinton et al. 1995]. This allows the memory architecture to recreate the language instruction when performing recognition and the motor directions when instructed to perform a behaviour.

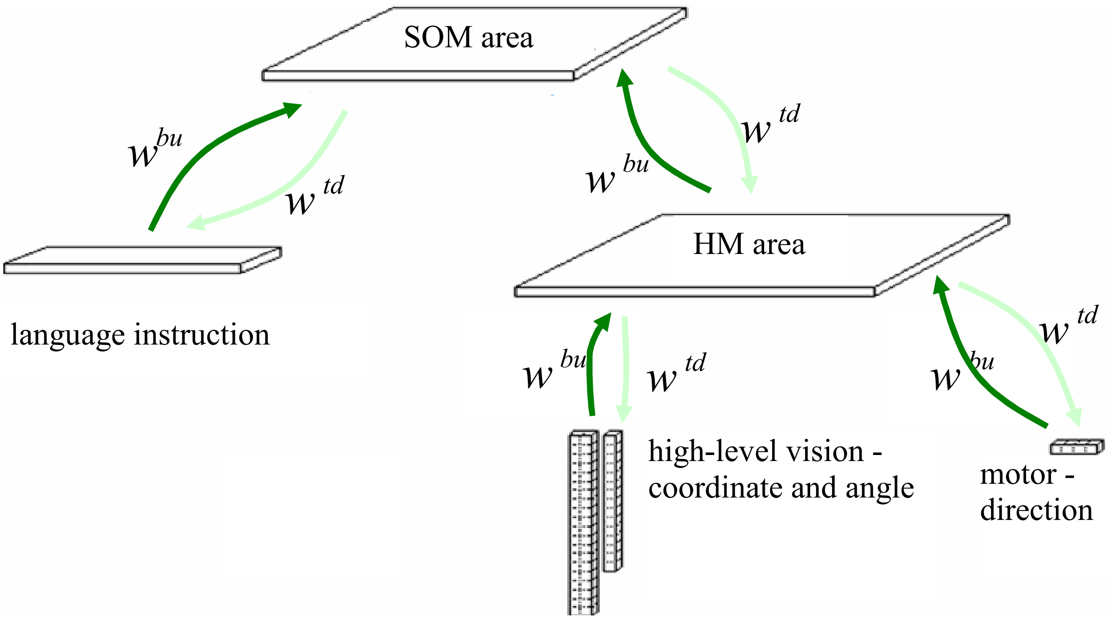


Figure 4.1.2.2 The hierarchical memory architecture for grounding language based on mirror neuron system (After Wermter et al. (2005)).

In this hierarchical memory based architecture there is the association of the motor and high-level vision inputs using the first hidden layer using the Helmholtz machine learning algorithm shown on the diagram as the HM area. The activations of the first hidden layer are then associated with the language instruction region input at the second hidden layer based on the self-organising map (SOM) learning algorithm, shown as SOM area. Using a Kohonen self-organising map in this architecture allows the features produced on the Helmholtz machine hidden layer to relate a specific motor direction with the appropriate high-level visual information for each behaviour based on the language instruction. The language instruction input is omitted when the student robot architecture is required to take the other inputs that are gained from observing the teacher robot and recognise the behaviour that is performed. When the motor direction input is omitted the student robot is required to perform a behaviour based on a language instruction. The architecture then continuously receives its own current x and y coordinates and angle and the language instruction of the behaviour to be performed.

Without the motor direction input the grounding memory architectures has to produce the appropriate motor activations which it has learnt/memorised from observing the teacher to produce the required behaviour. Recognition is tested by comparing the units which are activated on the language instruction area with the activation pattern belonging to the verbal instruction of the appropriate behaviour at each time step. Behaviour production is tested by comparing the performance of the behaviour at each time step by the student robot with the teacher robot.

4.2 Neurocognitive evidence on word memory representation

Neurocognitive evidence of Pulvermüller and his colleagues at Cambridge [Hauk and Pulvermüller 2004, Pulvermüller 2001, Pulvermüller 2002, Pulvermüller et al. 2001, Pulvermüller et al. 1999, Pulvermüller et al. 1999] can offer inspiration on how the brain represents and memorizes words and how this can be incorporated in the language emergent behaviour of the intelligent agent in the ACORNS projects. This evidence indicates that words are represented/memorised and processed using Hebbian learning, synfire chains and based on semantics features. Hebbian learning is the basis of higher cognitive behaviour through a simple synaptic approach based on cell assemblies for cortical processing and so can form the basis of some of the learning in the memory architecture of the ACORNS architecture.

Synfire chains are formed from the spatiotemporal firing patterns of different associated cell assemblies and rely on the activation of one or more cell assemblies to activate the next assembly in the chain [Pulvermüller 1999]. The word represented by a synfire chain depends on which cell assemblies are activated and when this occurs [Pulvermüller 1999, Pulvermüller 2003]. Once a spoken utterance is received the appropriate cell assembly representation should be ignited due to the strong feedforward and feedback structure. Hence, neurocognitive evidence on word representation and processing in the cerebral cortex suggests that cognitive representations are distributed among cortical neuronal populations [Pulvermüller 1999, Pulvermüller 2002, Pulvermüller et al. 1999]. Word meaning is critical for determining the cortical populations that are activated for the cognitive representation task.

When considering of the cell assemblies that process and represent content words, Pulvermüller (1999) states that activation is found in both hemispheres of the cerebral cortex. Semantic word categories elicit different activity patterns in the fronto-central areas of the cortex, in the areas where body actions are known to be processed [Hauk and Pulvermüller 2004, Shtyrov et al. 2004]. Perception words are represented by assemblies in the perisylvian cortex and posterior cortex [Pulvermüller 1999, Pulvermüller et al. 2001] and nouns related to animals activate the inferior temporal or occipital cortices [Pulvermüller 1999, Pulvermüller 2001, Pulvermüller 2002]. Emotional abstract words activate the amygdala and cells in the limbic system more than words associated with tools and their manipulation [Perani 1999]. The link between the assemblies in these two regions is achieved through the amygdala and frontal septum [Pulvermüller 1999]. However, function words that have a grammatical role are limited to the perisylvian cortex. For action words that involved moving one's own body the perisylvian cell assembly is also associated with assemblies in the motor, premotor and prefrontal cortices [Pulvermüller 1999, Pulvermüller 2003]. Assemblies that depict vision words are found in the perisylvian and visual cortices in parietal, temporal and/or occipital lobes.

The importance of semantic features compared with lexical categories in the grounding of language according to Pulvermüller (2003) is shown by the research of Kiefer (2001). Kiefer (2001) found when a set of nouns and verbs are tool- and animal-related words the division of cerebral cortex responses is based on whether they were tool or animal words. In a similar way, Perani et al. (1999) observes that although words in a sample set are nouns and verbs, the differentiation in the cell assemblies that represented them is not based on this but whether the words related to manipulating an object or are more abstract words. Pulvermüller [Pulvermüller 1999, Pulvermüller 2002, Pulvermüller 2003] argue it is important to relate the neurons that represent the word form with those neurons associated with perception and actions that reflect the semantic information on a word.

For content words the semantic factors that influence the cell assemblies come from various modalities and include the complexity of activity performed, facial expression or sound, the type and number of muscles involved, the colour of the stimulus, the object complexity, movement involved, the tool used and whether the person could see herself doing this activity [Pulvermüller 1999, Pulvermüller 2003]. The combination of these characteristics into a single depiction is produced by pathways linking sensory information from diverse modalities to the same neurons. For objects the semantic features represented by cell assemblies typically relate to their colour, smell or shape. If a word is repeatedly presented with a stimulus the depiction of this stimulus is incorporated into the one for the word to produce a new semantic feature. Fundamentally, words are depicted via regions historically known as language regions and additional regions connected with the word's semantic features. Moreover, there is evidence for distributed cortical assemblies that bind acoustic, visual and motor information and stressed the role of fronto-central premotor cortex as a prominent binding site for creating neural representations at an abstract semantic level [Pulvermüller 2003].

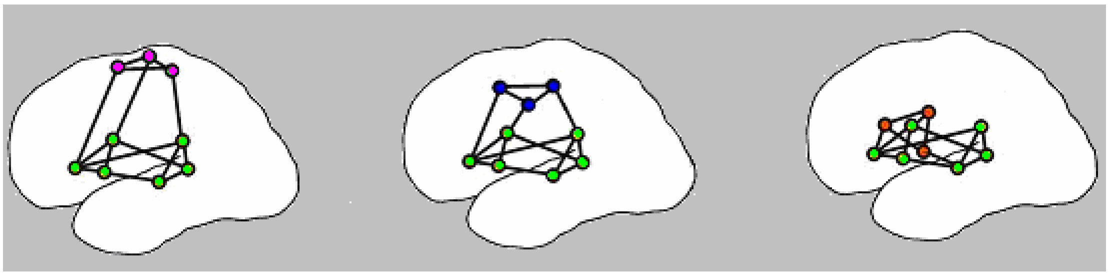
4.2.1 Neurocognitive evidence of action verb representation and processing

The neurocognitive evidence of Pulvermüller on action verb processing provides the basis for grounding of action verbs in actions. As well as a division between categories based on whether a word is action related or not [Pulvermüller et al. 1999], Pulvermüller stated that there is finer grained grounding of language instruction in actions. This creates a division of representation in the cerebral cortex based on the part of the body that performs that action between leg, head and hand [Hauk and Pulvermüller 2003, Hauk and Pulvermüller 2004, Pulvermüller 1999, Pulvermüller 2002, Pulvermüller 2003]. It is well known that there is a division in the motor cortex between the regions that performed head/face, hand/arm and leg actions [Penfield and Rasmussen 1950]. For instance, the region of the motor cortex that controls face movement is found in the inferior precentral gyrus, hand and arm in the medial region of the precentral gyrus and the leg actions are located in the dorsomedial area [Pulvermüller 2002, Pulvermüller 2003] [See Figure 3.2.1.1]. Given the difference in the regions of the cortex that are responsible for performing actions it is also believed by Pulvermüller that a similar difference can be identified when representing action verbs and so grounding language instruction in actions based on the part of the body that performs the action [Pulvermüller 2002].

Pulvermüller and his colleagues performed various experiments [Hauk and Pulvermüller 2003, Hauk and Pulvermüller 2004, Pulvermüller 2001, Pulvermüller 2002, Shtyrov et al. 2004] on cerebral cortex processing of action verbs to test their hypothesis on the representation of action verbs. These included experiments where (i) different groups of subjects are given leg-, arm- and face-related action verbs and pseudo-words and all asked to state whether they are a word; (ii) subjects were asked to use a rating system to answer questions on the cognitive processes a word arouses; (iii) subjects rank words based on whether they are leg-, arm- or head-related; and (iv) there is a comparison between hearing walk- and talk-type verbs. In these experiments EEG electrodes are positioned at various points along the scalp to produce recordings of cerebral cortex activation. From these experiments areas are identified where the activation is the same for all action verbs and more importantly are different based on the body part the action verbs relate to.

The regions these experiments identify as the same for the three types of action verbs are left hemispheric inferior temporal and inferior frontal gyrus foci. Turning to the differences between the three types of action verbs, there is greater activation for face words according to Pulvermüller and his colleagues in the frontal-lateral regions of the left hemisphere close to the premotor cortex associated with face and head. For face and leg related action verbs there are different regions along the motor strip that are identified to process verbs from these two verb categories. Leg-type words produce greater activation in the cortical region used to produce leg actions and for the face words there is greater activation in the inferior regions near to the face region of the motor cortex [Pulvermüller et al. 1999]. It is found that hand-related words are located in more lateral regions of the cortex than leg words. Hence, consistent with the somatotopy of the motor and premotor cortex [Penfield and Rasmussen 1950], leg words elicited greater activation in the central cerebral cortex region around the vertex, with face words activating the inferior frontal areas, thereby suggesting that the relevant body part representations are differentially activated when action words are being comprehended.

In addition, the average response times for lexical decisions is faster for face-associated words than for arm-associated words, and the arm-associated words are faster than leg ones. When considering the findings at 220ms, leg words had greater activation in central regions beneath Cz and C1 compared with face words. There is also greater activation in the right parieto-occipital areas for arm words and leg words relative to head words. The location of cell assemblies that Pulvermüller identified to process/memorise action verbs based on the body part they relate to can be seen in Figure 3.2.1.1. In this figure it is seen that the three types of action verb share certain cell assemblies (yellow circles) but differ on others which are shown green for the leg-related words, blue for the arm ones and red for the face-related words.



leg-related words arm-related words face-related words

Figure 3.2.1.1 The cell assemblies found to be associated with the processing of action words based on body parts (After Pulvermüller 1999).

It is also found that the cerebral cortex regions that performed a specific action are also part of the associated language representation. Words associated with hands activated the hand motor cortex, words associated to the leg activated the leg motor cortex and words associated with face activated the face motor cortex [Pulvermüller et al. 2004]. Hence, some of semantic features and, therefore, part of the meaning of the action verbs are represented through activation along the motor strip and in the premotor cortex. The evidence of the experiments performed by Pulvermüller and his colleagues [Hauk and Pulvermüller 2003, Hauk and Pulvermüller 2004, Pulvermüller 2001, Pulvermüller 2002, Shtyrov et al. 2004] point to word semantic features being represented in different parts of the cerebral cortex in a systematic way. Particularly the representation/memorising of the word is related to the actual motor and premotor regions of the cerebral cortex that performed the action. Furthermore, mental imagery of the action without actual body movements leads to specific activation in the fronto-central premotor cortex [Buccino et al. 2001] pointed to the role of regions associated with producing the action in language representations. These semantic representations need to be thought of as being topologically specific, that is, networks represented concepts with different meanings may have different cortical distributions.

4.2.2 Computational self-organising language memory

The unimodal memory architecture developed Wermter et al. (2003) explores the use of semantic features as action descriptions in a self-organising memory to achieve grounding of language instruction in actions [Figure 4.2.2.1]. The unimodal architecture uses a hierarchical approach so that a complex problem is divided into less complex recognition activities. This architecture undertakes a more coarse classification at the lower level and a more precise classification at the higher level, with the output from the coarse level controlling the input to the finer classification. The architecture contains a self-organising map (SOM) to perform the coarse clustering that related the action verb representations with the appropriate body part by clustering the verbs in different regions. The topological representation of the SOM fits well with the neurocognitive evidence on the action verb representation principle which pointed to a topological representation, with the cell assemblies that represent a word distributed across the motor and premotor cortex [Pulvermüller 1999, Pulvermüller 2003]. At the next processing level of the unimodal grounding architecture, there is finer clustering through a SOM for each body part taking the semantic features verb representations identified at the coarse level to be associated with the appropriate body part to cluster the actual action verbs.

This memory architecture is able to recreate, at an abstract level, the action verb representation and processing principle by clustering actions based on the body part used. The unimodal architecture is able to use the concept of semantic features by using the actual sensor readings to act as these features. Hence, an action verb such as 'put' consisted of an interrelated set of features that provides an indication of the robot internal state. The architecture went beyond the neurocognitive evidence of Pulvermüller on action verbs by also identifying the actual action verbs.

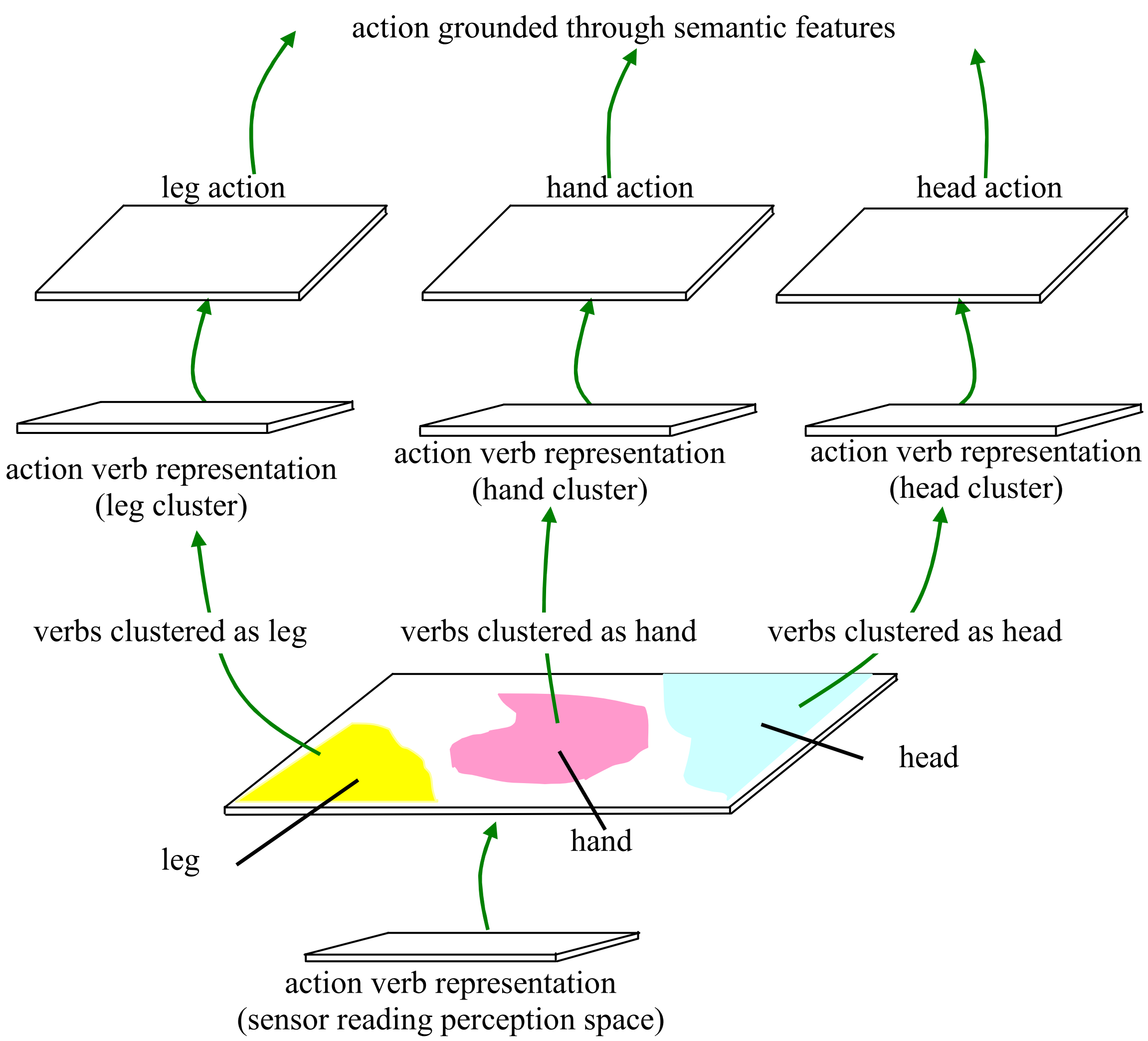


Figure 4.2.2.1 Unimodal memory architecture based on modular distributed and self-organising memory (redraw from Wermter (2003)).

4.3 Learned acquisition of language as a description of action

Vogt (2000) develops a language acquisition memory model for describing actions using a coordination game, with one robot following another while they developed a lexicon about the actions performed. In Vogt's (2000) model, the robots ground symbolic meaning in the real world based on non-symbolic sensorimotor data. The role of the robots is to produce categories and a lexicon to communicate actions such as 'go right'. The robots did a set of 'follow me games', with each experiment split into two parts: a development element and a testing element.

The teacher robot performs the action and communicated what it is doing based on the word form that is most associated with the action. In response, the student segments and categorises the teacher's actions and tries to comprehend what it is saying. The approach used for categorisation of the action segments is similar to the one of Rosenstein and Cohen (1998). After categorising the time span containing a motor actions the student tries to comprehend the word form it received by comparing its lexicon with the word form. If the matching memorised word form is found, its related meaning is matched with a categorised meaning of the actions and the lexicon is produced. The student matches the categories linked with action segments in the time frame to the related word meaning. The action game is completed successfully if the student knows the action performed by the teacher, has the correct word form and lexicon and is positioned behind the teacher. In the approach of Vogt (2000), the teacher has to produce a new word form if it has not already memorised one associated with an action. In response, the student does not have a meaning related with this word form and so related the action with the word form provided by the teacher. Where an action segment is recognised by the student and the word form is known, the memorised relation between that word form and the action is strengthened while the relation with other actions was reduced.

The development phase involves gaining the categories and a lexicon about the categories when the student is following the teacher. The success of the system is measured in terms of identified and communication success. Identified success determined the average success of the categorisation as a moving average over the last 50 language games and communication success measured the average success over the following 50 language games. Identification success grew very quickly to 90% with categories falling into 6 main action groups: 'backward', 'forward', 'go-left', 'go-right', 'go-left go-right' and 'go-right go-left'. The communication success in the development stage grew to 40% over the first 1000 language games and up to 55% after 10000 language games. The test stage determines whether or not the robot can do the following game where only the lexicon is used. The student attempts to comprehend the word form by producing a related action category association and executing the action. In the test stage the communication success is around 50%.

4.4 Computational memory models of grounding of words in actions based on meaning

The grounding of words in actions based on meaning approaches is examined by considering the approaches of Bailey (1995) and (1997) and Bailey et al. (1998) and Roy (2004). Bailey (1995) and (1997) and Bailey et al. (1998) grounds action verbs in motor actions by using an x-schema to represent actions by extending Petri nets with places represented as circles and transitions as rectangles. Places had tokens and held predicates on the world state and internal state, and transitions are the active elements. When all the places pointing into a transition had a token, the transition is activated and a new set of tokens are positioned at the output positions. Such x-schema are used to represent an action such as to slide an object on a table.

The model also contained a memory arrangement know as a f-struct to ground language in actions using static features that are given a value and so acted as the semantic features of the action verb. A special f-struct acts as a link between the word and the action with a bi-directional connection with the x-schema that translates actions into semantic features. A critical feature of the f-struct is the name of the x-schema of the action performed as well as motor features such as force, speed, direction and the position of the robot hand. Each sense of a verb like 'push' or 'pull' is represented as an f-struct using probability values, with the most appropriate meaning selected based on the world state. A significant benefit of the memory model is the ability to cope with multiple word inputs and multiple senses of an action. The model learns/memorises examples of verb word/action pairs and employs Bayesian Model Merging to deal with different verb meanings where representations of prototypical motor-actions for a verb are created or combined using a minimum length description condition for clustering. Using this approach, if the f-struct features for a new action verb representation are similar enough to previous examples it is merged into a single description of the action verb, however if the f-struct is sufficiently different a new sense/memory of the action verb is created.

Another approach for grounding language in actions based on meaning was devised by Roy (2004) who concentrates on defining the structures and theoretical rules that are the basis of the systems. This approach does not allow the robot to self-learn these rules and simply offers an overall framework for modelling the behaviour associated with specific types of robots. The approach uses three criteria: (i) objects, characteristics, actions and circumstances are produced from the same primitives; (ii) information created from vision and language have a common depiction structure; and (iii) motor actions and language actions are stated so that the robot can decide by using language and motor actions how to achieve a goal. A robot has a belief that persists over a period of time and so could act as a prediction of the future. In this approach signs are represented as patterns in the world that indicated a particular situation and were seen as natural, intentional and indexical. In a robot the only way for a sign to influence the robot is through its sensors.

Roy (2004) develops a graphical notation to use with this approach for representing schemas. In the approach signs are fundamental, with a sign being patterns in the world that can be interpreted. Beliefs about these signs are produced in the approach by using a-signs and d-signs. A-signs are distributions of signs as they are shaped by incoming signs and produce forecasts about signs. The meaning of an a-sign to the robot is based on its role in guiding the robot's view, interpretation and control procedures and is represented in the schema as an oval. The second form of belief about signs is the d-sign related to the output of a discrete categorisation procedure that links a continuous domain a-sign to a discrete domain and is represented as a rectangle. The physical characteristics of the robot allow a set of primitive actions to be associated with it. Actions provide an action projection representative component whose notation is a diamond, with a success or failure output related to the action. Actions are seen to produce d-signs and so link actions to signs. Roy (2004) shows it is possible using his notation to construct structured schemas that depict complicated behaviours and so the basis for the grounding of language in actions.

Objects are produced using the representation based on incorporating action and property schemas. These objects are memorised in relation to their interaction with their environment and relationship between actions that acted upon them. For instance, the notation and the different features of the cup are used to represent that a cup can be lifted by a gripper and also recognised based on features such as its colour and shape by a camera. Speech is seen as a canonical intentional sign which involves data representations that have meaning and are either descriptive or directive. The term descriptive is used for a statement about the condition of the world, while the term directive was used for instruction that had the goal of causing a change.

Above we have described two memory based approaches that ground language in actions based on meaning that are likely to influence the memory architecture for the ACORNS project. Both these grounding of language in actions approaches provide a graphical notion for representing the action verbs. The approach of Bailey indicated the importance of semantic features in the representation of action verbs.

4.5 Memory models of grounding of language in objects

The form that memory grounding of language in objects has taken will be investigated below as the emergence of language in the ACORNS intelligent memory agent incorporates interaction with objects in their world. As noted by Ziemke (1997) a great deal of research into grounding takes the form of the association of language and objects. This involves the development of a model that matches symbols with sensor information related to the objects [Coradeschi and Saffiotti 2003].

Roy (2001) and Roy and Pentland (2002) developed CELL to ground language in objects using a robot equipped with a camera and microphone. Although this approach made use of intelligent learning it does not use neuroscience evidence at a significant level. However, the robot is able to ground language by associating and memorising the symbolic representation of language (utterance) with semantic features of the utterance (visual representation). CELL represents/memorises 3-dimensional objects by combining multiple 2-dimensional views, which are depicted by histograms of local characteristics from the 2-dimensional representations. By producing multidimensional histograms of the objects it is possible to compare the objects based on statistical functions.

The CELL model addresses the problem of identifying words from speech and attaching meaning to those words. Inputs to the system are spoken utterances and views of objects, in a manner which matches how children learn this association when receiving instruction by a parent. The speech data consists of actual interactions between a child and the parent with the problems typically associated with actual speech recognition. Short-term memory is used to store utterance-shape pairs and produce hypotheses about the association by extracting part of the utterance and linking it with the observed shape. This association is placed in long-term memory which is consolidated over various observations. Spoken utterances are depicted as sets of phoneme probabilities, which are converted into a spectral depiction using the Relative Spectral-Perceptual Linear Prediction (RASTA-PLP) algorithm. A recurrent neural network examines RASTA-PLP coefficients to predict phoneme and speech/silence probabilities. To identify phoneme boundaries the recurrent neural network outputs are seen as state emission probabilities in a Hidden Markov Model structure.

Word learning/memorising occurs through two levels of memory, short-term and long-term memory. Input depictions of speech paired with visual objects are temporarily kept in the short-term memory. Every entry in the short-term memory includes a phoneme probability vector depicting multiword speech, and a group of histograms of the object. A recurrent filter acts on the short-term memory by searching for repeating subsequences that happen in similar visual contexts. The output of the recurrent filter is the first attempt to split speech into words. The long-term memory included two kinds of data structures. The first is AV-prototypes which depict hypotheses of likely words of the target language. The second kind of data structures in long-term memory are lexical items that are produced by consolidating AV-prototypes. This procedure determines clusters of AV-prototypes which could be combined to model consistent intermodal patterns between multiple observations.

The study by Roy (2001) and Roy and Pentland (2002)consists of parents interacting with their child in a natural manner while playing with objects. For each set of speech data, the CELL model is used to gain the 125 highest scoring lexical elements, which are examined based on segmentation accuracy, word discovery and semantic accuracy. CELL is found to outperform an acoustic-only model on segmentation accuracy, word discovery and semantic accuracy. The acoustic-only model struggles on the segmentation accuracy test: only identifying correctly the word boundaries 7% of the time compared with 28% for CELL. On word discovery for CELL 72% of the lexical elements are single words, with the acoustic-only model's performance being 31%. For the final measure, semantic accuracy, there is the greatest difference between CELL and the acoustic-only model (CELL 57% and acoustic-only 13%).

An additional approach to memory based ground language in objects is taken by Steels and Kaplan (2001) who use the Sony AIBO dog robot to learn in the manner of a 12 month old child. Experiments are performed based on the naming of three objects: a red ball, a yellow puppet known as Smiley and a small ABIO type figure called Poo-Chi. In these experiments the level of social mediation is varied, with the aim being that the robot dog grounds language in the visual images.

The robot includes a script consisting of schemas to carry out the classification game. In the first experiment, once the human believes that they have achieved the attention of the robot it was shown the object and the human says its name. The robot repeats the word in order to ensure that it has heard the word correctly and the human confirmed this is the case. By ensuring that the interaction between the robot and the human is tightly coupled, the human mediator assists the learner robot to achieve social learning by offering feedback. A simple interaction between the human mediator and the robot can be seen in Table 4.5.1. The robot associative memory links object views and words. Word learning occurs through reinforcement learning where the human offered positive feedback. The robot classifies the object and then received feedback whether it has correctly identified the object. To perform object classification diverse 'views' of the object are stored and the nearest neighbour approach is used.

Table 4.5.1 Example interaction between mediator and AIBO.

Human	What is it?
AIBO	Poo-chi
Human	No; listen; Smiley
AIBO	Smiley?
Human:	Yes

In the second experiment to consider the role of interaction in grounding of language in objects, Steels and Kaplan (2001) uses a database of interactions between a human and robot that contains social learning dialogues to train the network. Unsupervised clustering is performed by the

Expectation-Maximization algorithm that identified 8 clusters, but these are not related to those required to learn the words. The clustering relates more to the light conditions and background than the actual objects. The clusters that are associated most to the objects are cluster C3 which relates to Poo-Chi, cluster C2 which relates to the red ball, and cluster C5 which relate to Smiley. This however only achieves a classification rate of 30%.

In the final experiment Steels and Kaplan (2001) study the impact of reducing the mediation between the robot and human. The robot moves around freely where there are three objects: Poo-Chi, Ball and Smiley. When the human sees the robot looking at an object they state the name. However, this includes the problem that the human might not know exactly what the robot is looking at and so provides a description for the wrong object. Using a dataset of 150 images (50 for each object) this approach achieved a classification rate of 59%. When considering the errors it is found that there is mainly misclassification between the Poo-chi and Smiley. This research shows the importance of social learning in the process of grounding language in objects, which is incorporated in the ACORNS project through the interaction between the .intelligent agent and the caregiver.

Mavridis and Roy (2005) developed a memory based model that is able to ground language in activities known as the grounded situation model. The memory model is located at the centre of a structure that has around it modules related to language, perception and action. The grounded situation model acts as a memory that stores current, present or imagined events and includes a representation of the knowledge that the robot has gained related to itself, the human it interacts with and objects in its environment. Layer 1 includes stochastic representations of characteristics that could be measured using sensors. Layer 2 includes representations of characteristics for use as action controllers. Layer 3 produces discrete categorised representations of characteristics for use in language interactions. The grounded situation model is based around a modular approach. The situation model is used to store and pass on the current state of the grounded situation model. Particular modules are used to recommend changes to the grounded situation model which are passed on by the situation model. The inquiry module allows the grounded situation model to answer simple requests and the rememberer module can recall past experiences to incorporate temporal relationships.

The grounded situation model is tested using the token test, which is a typical language acquisition test for children and as such involves placing objects on a table and given instructions. An instruction can be 'when I pick up the red tile you move the blue one'. The token test has 5 levels of difficulty of which the grounded situation model is designed to pass the first two. In addition to this it is also able to acquire an understanding of environment that it has only been described not actually seen using its imagination module. By use of a hierarchical object arrangement, three layer property depictions, and a recursive model approach it is felt that the grounded situation model is an approach to allowing robots to acquire the capacity to ground spoken language with the real world.

5 Reinforcement learning

The learning between the caregiver and the ACORNS learning agent is to incorporate a form of reinforcement learning memory, where by the caregiver offers feedback to the learning agent to influence its learned/memorised behaviour. Reinforcement learning is a field of machine learning where an agent can learn by trail and error to perform an action to achieve a reward or to stop a punishment [Daw and Doya 2006].

5.1 Biological basis of reinforcement learning

In the brain the area that has been most associated with reinforcement learning is the Basal Ganglia [Doya 1999]. The principal input nuclei of the basal ganglia is the stratium which gains inputs from the motor cortex, prefrontal cortex and the intralaminar nuclei. Graybiel (2005) states that there is a great degree of evidence that the basal ganglia is involved in the learning process of selecting the most appropriate action based on the current environment. A fundamental element of basal ganglia research has been that cortio-basal ganglia circuits to achieve sequential learning through a form of learning by trial and error. Using this form of trial and error actors examine their environment and altered their memorised behaviour using reinforcement signals to achieve the optimum given the situation [Graybiel 2005]. Neuroscience support exists that the basal ganglia uses a review-specific approach based on release of dopamine to learn behavioural approaches based on online feedback [Doya and Sejowski 1995].

The principal input nucleus of the basal ganglia is the striatum, which is felt to be linked to motor pathologies, learned habitual and dopamine based learning with activity in this area founded on actions what their impact will be [Daw and Doya 2006]. The basal ganglia is found to have various inhibitory pathways with the striatum receiving inputs from the cerebral cortex. Studies have shown that the basal ganglia is involved in reinforcement type learning and performance of goal-oriented sequential action memorisation [Doya 1999]. By using reward-predicting behaviour of dopamine neurons and the alteration of actions during learning the basal ganglia is able to achieve reinforcement learning. Interest in the basal ganglia as an area of reinforcement learning occurred as a result of the experiments performed by Wilfram Schultz who considered the dopaminergic neurons in primates when learning behavioural activities [Joel et al. 2002, Schultz and Dickinson 2000, Schultz et al. 2000].

Schultz et al. (2000) when considering the processes involved in reward and so reinforcement learning found the areas of the brain associated with reinforcement learning are the orbitofrontal cortex and the basal ganglia, whose neurons have different roles in rewarding actions. Neurons in the orbitofrontal cortex are associated with reward-predicting instructions, the period straight before the reward and after the rewards. The neurons in the striatum of the basal ganglia are associated with the expectation or identification of reward, with the performance of an action being related to the receiving of a reward at the end of an action. In experiments with primates, activation of neurons in the striatum are dependent on a liquid reward at the end of the trail, which indicates that the neurons in the striatum are very much dependent on the expectation of the reward associated with actions that create the reward. According to Schultz et al. (2000) dopamine neurons acted like a global reward signal which projects to neurons in the striatum and the frontal cortex. The neurons in the orbitofrontal cortex have a reward link in a selective manner to postsynaptic neurons. In such an approach a teaching role is performed for a restricted number of cells instead of having a global impact. The dopamine, orbitofrontal and striatum neurons respond to predicted rewards. Orbitfrontal and striatam are found to be excited by a predictable reward and dopamine cells inhibited by the omission of a reward. Striatum neuron activation depends on the performance of a movement that is likely to produce a reward. A type of reward process is found where there is activation in the orbitofrontal, striatal and dopamine neurons before a predicable reward occurred. This is likely to indicate that the primate has an internal representation of the reward it would receive if it performs a specific action. Schultz et al. (2000) in their experiments are able to identify that the expectation of a reward has an impact on task-associated activity of the striatum.

There are two opposing views regarding the association between the prefrontal cortex and basal ganglia. The first conventional view is that the prefrontal cortex drives the learning of the basal ganglia [Laubach 2005]. While the second view as stated by Graybiel (2005) is that as well as the dopamine system 'teaching' the striatum, the basal ganglia teaches the cortex through a cortico-basal ganglia loop through the straito-pallido-thalammocortical pathways. There is support for the view that the basal ganglia instructs the cortex with Pasupathy and Miller (2005) noting that learning associated changes happen much earlier in the striatum than the cortex. They found for associative learning the striatum has a quicker change in neural activity compared with the slower prefrontal cortex which relates to a slower improvements in behavioural.

Pasupathy and Miller (2005) examined the time course of learning associated alternations in the dorsolateral prefrontal cortex and the striatum. Primates are trained to look at a light for 0.5s and a complicated cue shown at the fixation point. Following a delay of 1s, the fixation point spot is removed and the target is presented in each eye. The primate gets a reward of juice if its eye movement is towards the correct target. The primate is able to determine the appropriate behaviour using trail and error by learning the relationship between the input and the reaction. Once the relationships are learned/memorised by the primate the input-response pairs are reversed and the primate is required to relearn/rememorise the appropriate behaviour to input. The striatum is found to have direction-specific firing almost straight away following the reversal of the input-reaction pairing. However, the prefrontal cortex only gains direction selectivity following 15 trails and so the striatum is responsible for the training of the prefrontal cortex.

5.2 Reinforcement memory learning through the actor-critic memory model

Reinforcement learning alters the behaviour of an agent with the aim being to be able to achieve the level of reward for the conditions faced [Barto, 1995]. It is a powerful approach to produce goal-related sequential learning in a learned manner. A model according to Khamassi et al. (2005) that relates to reinforcement memory learning in the basal ganglia is that of the actor-critic model. Such an approach offers a suitable method for learning in the ACORNS intelligent agent and the interaction between the agent and the caregiver. The actor-critic model comes from the field of incentive learning theory where a reward such as a shot of fruit juice to the primate produces conditioned learning [Dayan and Balleine 2002]. It is one of the most prominent models of the basal ganglia to relate the dopamine neurons behaviour with the temporal difference prediction error signal [Joel et al. 2002]. In this approach the actor network selects the next action based on maximising the weighted sum of rewards in the future by using a network known as the critic. The critic stores the value that indicates the degree that the state assists in getting the actor to the goal. The actor can be seen as the matrix area of the basal ganglia and the critic equates to the striosomes in the dorsal striatum of the basal ganglia. The critic produces a dopamine-like, reinforcement signals that allows it to forecast the reward during the task, and so assists the actor to learn to select appropriate behaviours during the task

[Khamassiet al 2005]. For Joel et al. (2002) the actor-critic model is able to determine the weighted sum of rewards in the future, by combining the inputs from sensors with the approach of the actor. Hence, the actor-critic memory model developments a temporal difference learning rule that associates the reward with the inputs associated with this reward from sensor signals [Khamassi et al. 2005].

5.3 Computational memory models of actor-critic

According to Joel et al. (2002) one of the first actor critic-models of reinforcement learning is that of Houke et al. (1995). In this model Houke et al. (1995) argue that striosomal modules acts as the critic while the matrix function performs as the actor. Striosomal modules comprise of striatal striosomes, subthalamic nucleus, and dopaminergic neurons in the substantia nigra pars compacta (SNc). In relation to the temporal difference equation the main reinforcement in the temporal difference equation is associated with the main reinforcement to dopamine cells, the prediction of the future reward is associated with the indirect excitatory input for dopamine cells and indirect inhibitory stimulus is related with the prediction at the previous time step at the earlier time step. Houke et al.'s critic fails to include an exact timing approach, but instead depended on a slow and consistent inhibition of dopamine neurons [Joel et al. 2002]. The actor in this model uses matrix modules, which included the striatal matrix, subthalamic nucleus, globus pallidus, thalamus, and frontal cortex and generates signals that produce action or depict plans that require other systems to produce command signals.

A memory model was developed by Weber et al. (2003) to perform robot docking to a table behaviour using an actor-critic reinforcement associator networks approach. The model trains the weights to determine what the visual input is and where it is in the visual field to enable the robot is able to approach the required target and finally, using reinforcement learning it trains the weights to produce the appropriate motor outputs (forward, backward, left or right) to dock with the target. The inputs to the reinforcement network in this docking examples is the perceive state of the robot which differs for each location of the target and the angle of the robot to create a state space representation. The weights from the state space to the critic and the motor units are learned in a reinforcement manner. A reward is provided if the robot successfully docks at the table, otherwise no reward is given if the robot fails to dock as a result of hitting the table for instance. A critic value is allotted to each state, with those states that lead quickly to the goal having a higher critic value by strengthening their connections to the critic unit. The weights to the motor units which have been activated simultaneously are also increased, if the corresponding action produces a better state.

In this memory model by Weber et al. (2003) the actor-critic reinforcement learning algorithm is performed in the following steps to train the robot to perform/memorise the docking behaviour.

(a) Determine the target $\vec{p}$ and the perceived state $\vec{f}$
(b) Determine the activation for the critic $c = \sum_j w_j^c \cdot f_j$
(c) Determine the probability $P(m_i = 1)$ that the motor unit i is active

$$P(m_i = 1) = \frac{e^{2a_i}}{\sum e_i^{2a_i^j}} \text{ , with } a_i = \sum w_j^m \cdot f_j$$

(d) Move the robot based on the motor unit selected

(e) Determine the target $\vec{p}'$ and the perceived state $\vec{f}'$

(f) Determine the activation for the critic $c' = \sum_j w_j^c \cdot f_j^{'}$

(g) Determine the reward 1 if robot reaches goal, -3 if robot hits table or goes outside visual field; 0 other outcomes

(h) Determine prediction error between R the reward and $(c - \gamma c^{'})$ the critic evaluation , $\delta = R - (c - \gamma c')$

(i) Update the weights of the critic $\Delta w_j^c \propto \delta \cdot f_j$

(j) For the only active motor unit i update weights $\Delta w_{ij}^m \propto \delta \cdot m_i \cdot f_j$

According to Joel et al. (2002) an additional actor-critic model of reinforcement memory learning is that of Suri and Schultz (1998, 1999) who built on the model by Barto (1995) through an neural model of the actor and updating of the temporal difference algorithm to include the timed depression of dopamine activity when the reward is not provided. The critic learning rule is altered so that only the weight for the stimulus depiction element that relates to the actual stimulus-reward interval is modified. The actor according to Joel et al. (2002) of Suri and Schultz (1998, 1999) model includes a layer of units of which each depicted a particular action. This layer is able to associate input-action-action relations using the prediction error signal of the critic. By using the new version of the critic it is possible for a basic simple actor network is sufficient to solve relatively complex behavioral tasks. Suri et al. (2001) extended the models described in Suri and Schultz (1998, 1999) by developing a model that is bases closer the actor on the basal ganglia-thalamocortical circuitry and incorporates more complicated interaction between the critic and the actor. For Joel et al. (2002) the approach ensures that only one action is chosen by using the connection between the striathum and the basal ganglia output nuclei and at the cortical level and the use of a winner-take-all approach. The critic gains sensory and reward details, as well as gaining details on the planned and actual behaviour to learn input-reward and behaviour-input relations. The model by Suri et al. (2001) can plan in that it has the capacity to produce relations chains and choose actions based on the predictions from these chains [Joel et al. 2002].

A model of the actor-critic model was also developed by Khamassi et al (2005) to control an artificial rat. The actor element of the model is based on the approach devised by Gurney et al (2001a, 2001b) and is made up of a set of parallel channels each one depicting an action. The channels contain two different pathways, the first is the 'selection' pathway for action selection using feedforward approach and the second is the 'control' pathway which is responsible for managing the selection process. A cortex-basal ganglia-thalamus loop in the model enables it to determine the degree that each channel is involved in the selection process. In the memory system of Khamassi et al (2005) the inputs to the actor are strengthens associated with a specific action determined using 12 sensory values, a bias and a persistence level. For every timestep the action that has the greatest strength is performed. The strength associated with action i is:

$$Strength_i(t) = \left[\sum_{j-1}^{13} \mathrm{var}_j(t) \cdot w_{i,j}(t) \right] + persist_i(t) \cdot w_{i,14}(t)$$

$\mathrm{var}_{13}(t) = 1,\ \forall t, w_{ij}(t)$ acts as the synaptic weights for each action i, the related strength with input j. The aim of the learning process is to establish the weights that provide the most effective performance.

Khamassi et al. (2005) uses a model that combines multiple critics, each acting as a single neural unit and responsible for a specific section of the problem space. Two main critic models are considered the first the AMC1 critic model uses a mixture of experts with a network being used to decide which expert to use for which area of the robot environment. The second AMC2 critic model uses a mixture of experts where experts are allocated by hand to one sub-zone from the 30 sub-zones in the environment based on the division of the environment using visual perception information.

In the AMC1 model since the critic uses N experts, each expert k determines its own likely reward at timestep t

$$p_k(t) = \sum w'_{k,j} \cdot (t)\mathrm{var}_j(t)$$

$w'_{k,j}(t)$ acts as the synapse weights for the expert k.

The overall forecast by the critic is a weighted sum of all the experts' forecasts:

$$p(t) = \sum cred_k(t) \cdot p_k(t)$$

$cred_k(t)$ is the credibility associated with the expert k at timestep t based on the findings of a gating network.

$$o_k(t) = \sum w''_{k,j} \cdot \mathrm{var}_j(t)$$

$w''_{k,j}$ acts as synaptic weights for the gating cell k.

The credibility of expert k is then determined based on functions of output $o_f(t)$:

$$cred_k(t) = \frac{o_k(t)}{\sum o_f(t)}$$

Every expert for a critic has it own reinforcement signal using its prediction error.

$$\hat{r}_k = r(t) + gP(t) - P_k(t-1)$$

The weights of every expert k are updates using the following:

$$w''_{k,j}(t) \leftarrow w''_{k,j}(t-1) + \eta \cdot \hat{r}_k(t) \cdot \mathrm{var}_j(t-1) \cdot h_k(t)$$

where $h_k(t)$ acts as the element from each expert k towards the global prediction error of the critic

$$h_k(t) = \frac{cred_k(t-1) \cdot corr_k(t)}{\sum cred_f(t-1) \cdot corr_f(t)}$$

$corr_k(t)$ is the 'correctness' associated with a expert k.

$$corr_k(t) = \exp\left(\frac{-\hat{r}_k(t)}{2\sigma^2}\right)$$

σ acts as scaling parameter that is based on mean error of the experts.

The gating network weights are updated based on the following:

$$w''_{k,j}(t) = \leftarrow w''_{k,j}(t-1) + m \cdot diff(t) \cdot \mathrm{var}_j(t-1)$$

$$diff(t) = h_k(t) - cred_k(t-1)$$

m acts as learning for the gating network.

Model AMC2 is different from Model AMC1 in terms of how the credibility of the experts is determined. The aim of the AMC2 model is to remove the association of the credibility of the experts from their performance, by allocating a zone to the experts for the period of the study. The experts then learn to improve their performance as the study continues.

5.4 Reinforcement based dialogue systems

In response to the great deal of time and effort required for the hand-crafted development of language dialogue systems there is a movement towards using machine learning techniques and in particular reinforcement memory learning [Prommee et al. 2006]. Reinforcement learning has proved suitable for language dialogue systems where there is a transition from one state to the next. Language dialogue systems require moving from one representation state to the next, which can be achieved successfully through reinforcement memory based learning (Sutton and Barto 1998, Singh et al 2002). In such reinforcement learning approaches there is a state space, a set of actions, a transition function and a reinforcement reward function. The actor selects a particular action and moves into a new state using a transition probability distribution, based on which it receives a reinforcement reward. Singh et al 2002 uses an approach that combines reinforcement learning with Markov decision processes to produce a memory based dialogue management system. The aim of a dialogue system is to take a set of states and associate them with the appropriate group of actions. Usually in such systems a developer uses expert knowledge to produce these associations, however by using reinforcement learning it is possible to learn/associate them. This requires the development of a Markov decision process by using a corpus of dialogue interactions to establish a transition probability from one state to another based on a specific action by looking at the frequency of this occurring in the corpus. In the Markov decision process cumulative reward of an action given a state is the Q-value of the following states.

Pommer et al. (2006) have also used the reinforcement based Markov decision process to model a dialogue system based on human interaction with a robot. In this approach user and error models are collected through interactions between the user and a hidden human taking the role of the dialogue system. The user model is based on bigrams where the user's action is completely reliant on the previous system action. The ASR error model employed is based on prediction probabilities for a group of prediction activities. There is also error models produced for various multimodal features including gesture recognition. The robot acts as an early stage bartender, where the human is required to select an item from a group of items and the robot interacts with the human to establish the object and serve them with the item. The Markov decision process contains eight actions based on the interaction with human subjects. Eight state characteristics are used including information slots such as colour, location and object kind; how many candidates there is; the speech condition; ASR history, conditions for gesture tracking and last system action.

6 Working memory

Working memory according to Courtney et al. (1998) is the keeping of a limited amount of information in memory for a short period of time so it is available for use. Hence, working memory allows the person to maintain an active depiction of the conscious state for a short time period [Courtney et al. 1998]. It is the approached used to keep task specific information during the performance of a cognitive activity [Baddeley and Hitch, 1974, Priti and Miyake 1999]. Working memory is seen as process related and often seen as the 'blackboard' of the mind where the manipulation and storage of task related information occurred [Priti and Miyake 1999]. Although there are many regions of brain associated with working memory the ones that are seen as the most critical are the prefrontal cortex and various multimodal regions [Courtney et al. 1998, Brenchmann et al. 2007]. Brain-imaging studies showed that during a task involving working memory the prefrontal cortex and inferotemporal cortex keep a depiction of the visual input even when it is removed from view [Courtney et al. 1998]. Certain neurons in these regions are found to have a selective response to specific stimui [Mori and Horiguchi 2004].

6.1 Working memory models

Baddeley (1992) developed a model of working memory which he describes as a system in the brain that temporary stores and manipulate information so that cognitive activities such as language learning and understanding and reasoning can occur. Baddeley (1992) notes in the model that working memory is split into three main subsystems (i) the central executive that performs as an attention-control system; (ii) the visuospatial sketch pad for the manipulation of visual inputs; and (iii) the phonological loop which is used in the storage and rehearsal of speech based knowledge and the acquisition of languages. Baddeley (2003a) notes that the phonological loop includes the articulatory rehearsal procedure and can store a memory trace for a few seconds. The phonological loop has a role in the language acquisition process, which is supported in the finding that if the phological loop is disrupted by for instance word length it is not possible for subjects to learn a new language [Papagno et al. 1992]. As can be seen from Figure 6.1.1 the model is extended by Baddeley to include an episodic buffer that stores information from various sources of diverse modalities in the form of an episode [Baddeley 2003b].

An additional model of working memory is that of Cowan (1999) which is known as the embedded processing model of working memory [Mizuno 2005]. This model includes various subcomponents such as the central executive; long-term memory; active memory and attention focus. By using these elements it ensures all the information required for the task is available from one of these four elements. This model is described by Cowan as the 'virtual' short-term memory with the memory being active for up to 10 seconds and attention restricted to four separate items [Mizuno 2005, Cowan 1999, Cowan 2001]. As the focus of attention is limited in size if the information exceeds the memory size earlier items are likely to be removed [Mizuno 2005]. The activated component of the Cowan's model matches the passive stores in Baddeley model and the focus of attention is similar to the storage capability of the central executive in Baddeley model [Mizuno 2005]. In Cowan's model retrieval means placing items in the focus of attention, retrieval of information from long-term memory is restricted by the time frame of the task and retrieval from active memory is required has the memory will disappear in approximately 30 seconds [Mizuno 2005].

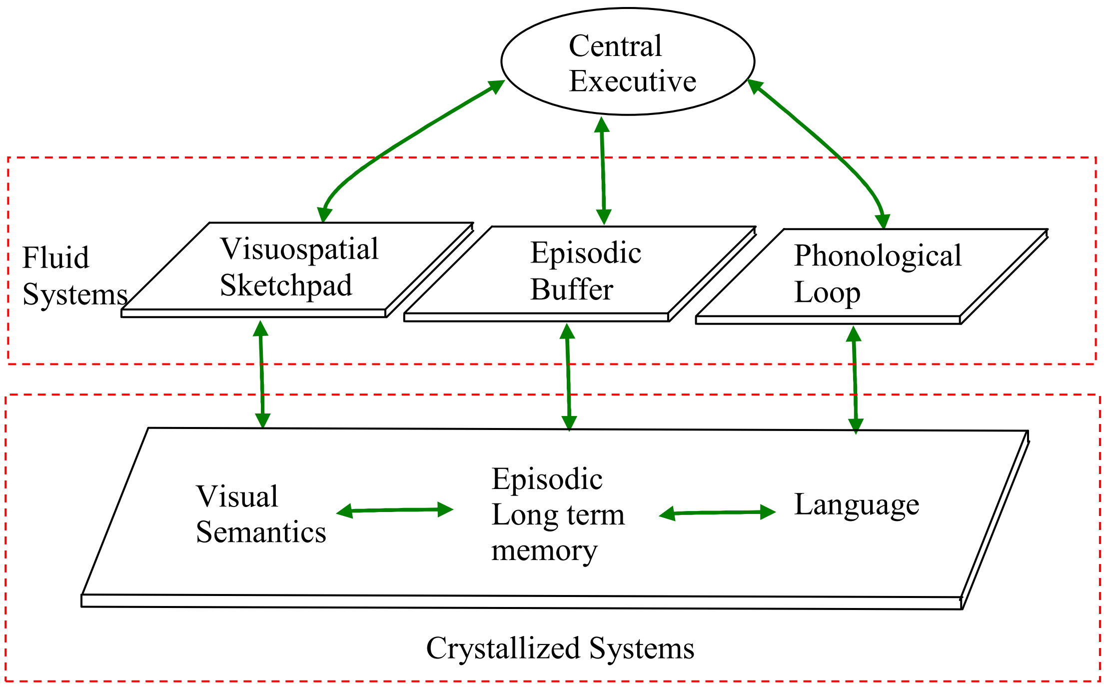


Figure 6.1.1 Diagram of Baddeley model of working memory model (After Baddeley (2001))

6.2 Computational models of working memory

One neural network approach that acts like a basic working memory model is known as the recurrent network. The simple recurrent network developed by Elman (1990) feeds back the activations in the state/hidden layer for the previous time step. Hence, the hidden layer does not only receive the external input but also this feedback [Bodén 2002]. By the use of a working memory type structure simple recurrent networks are able to create states that incorporate temporal relationships [Bodén 2002]. As can be seen in Figure 6.2.1 units in the previous state layer receives a copy of the activation values from the same unit in the hidden layer. The hidden layer units also receive inputs from units in the previous state layer as well as the external input units. The previous state layer enables the output values of the hidden layer to be kept in a form of short-term memory and then incorporated in the learning process by providing the previous activations of the hidden layer [Garfield and Wermter 2006]. Another type of recurrent network that has been developed is the Jordan recurrent network [Jordan 1986], which unlike the simple recurrent network does not feedback the previous activations of the hidden layer but feeds back the output activations to the hidden layer [Figure 6.2.2].

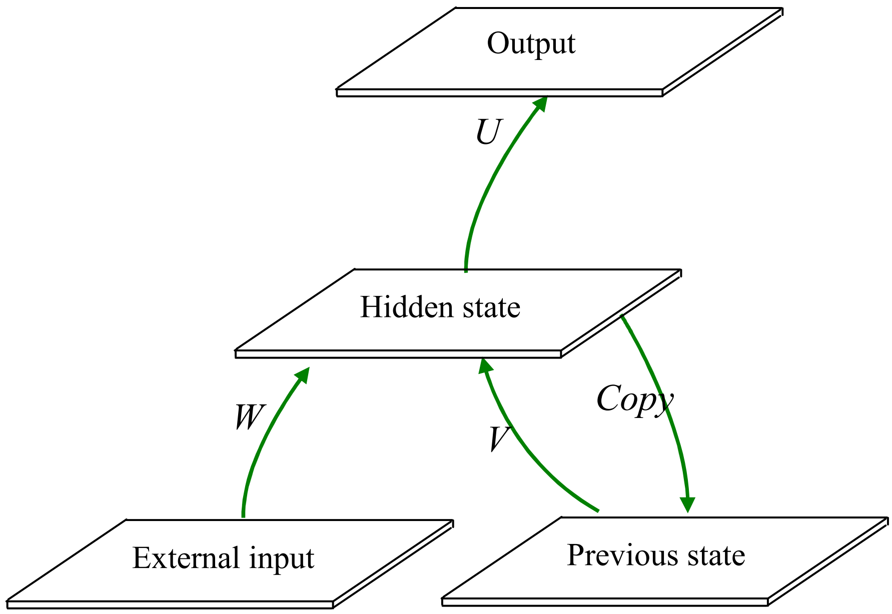


Figure 6.2.1 Simple recurrent network

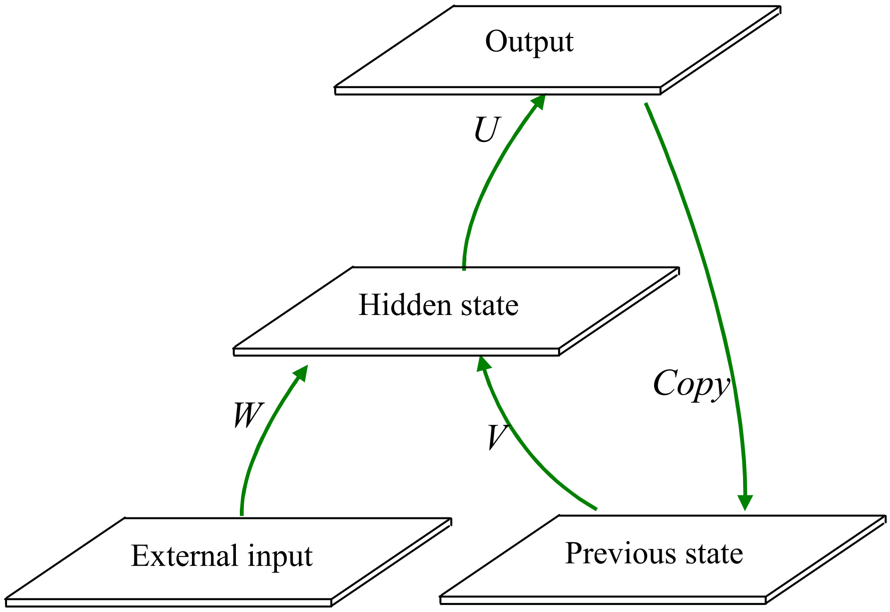


Figure 6.2.2 Jordan recurrent network

Due to capacity for recurrent networks to include working memory-like characteristics such as temporal processing they have been used a great deal for language processing. Garfield and Wermter (2002) use a simple recurrent network to perform operator assistance on a corpus of spoken language to telephone operators in order to position them in the appropriate call classes. The words in the calls are represented using a frequency measure based on the frequency the words appear in the specific call class. The input to the network is one word at a time, where it receives the frequency value for that word in each class. The network creates a value to indicate the class that the utterance was positioned in. Once the utterance is completed the previous state layer is re-initialised to zero. An utterance is allocated to specific class if at the end of the sequence it has a value greater than 0.5. The simple recurrent network when classifying unseen samples of the utterances into one of seventeen classes is able to achieve a recall rate of 75% and a precision score of 85%. When compared with a feed forward network that did not have a form of working memory the simple recurrent network performed better which indicates that the information included in the memory does aid classification.

Wu et al. (1993) also used recurrent neural networks to incorporate a basic form of working memory when performing the task of classifying the language being spoken. The system was first developed using a feed-forward network but due to the capacity of the simple recurrent network to learn the associations in temporal data a SRN was used. French and English speech is taken from two people one female and one male which produced 100 seconds of speech. A simple recurrent network is used to classify the speech into either French or English. The network has 192 input units, 11 units in the state and the previous state layers and at each time step 400 ms of overlapping speech with each sample being 6.25 seconds of speech introduced using a time window which moved along at intervals of 25ms. To perform classification on the test data for 400ms section a decision is made to whether the speech is English or French from this it is possible to achieve 99.5% accuracy after 41 votes of a sample that is 1.75 seconds long.

An additional model of working memory devised by Omori and Horiguchi (2004) using Hodgkins-Huxley type neurons. Their memory model is made up of excitatory and inhibitory neurons, with some of the excitatory neurons being selective to a stimulus and other being non-selective. The network is fully connected using dynamic synapses [Figure 6.2.3] with 80 excitatory units of which 32 selective and 48 non-selective units; and there are 20 inhibitory units. The model is successful in reproducing the behaviour of working memory by subpopulation units continuing to spike despite the external input being removed and so the maintenance of memory [Omori and Horiguchi 2004]. It also exhibits the behaviour that if a subpopulation is displaying persistent firing behaviour and the external input is enhanced to another subpopulation the persistent firing will either continue or be altered based on the level of the external input and the strength associated with the NDMA synapses.

A model of visual working memory was developed by van der Voort van der Kleij et al. (2003) in the ventral prefrontal cortex which contains a blackboard depiction of all objects in memory in order to bind their features together. When too many objects are in working memory and so are placed on the blackboard they start to interfere with one another. The posterior infero-temporal cortex links to a layer of the ventral prefrontal cortex. The representation in the posterior infero-temporal cortex and a layer of the ventral prefrontal cortex include combined depictions of the location and characteristics of the object. A further layer of the ventral prefrontal cortex links to higher regions to the visual cortex that processes and depicts the location and object features information. These links give feedback, with the lower layer of the ventral prefrontal cortex having a distributed representation of the distinctive features of the object. The lower and the top layers of the ventral prefrontal cortex interact using a feedback-feedforward approach, which produces activations in a third layer where activations are close between the lower and upper layer of the ventral prefrontal cortex.

The blackboard arrangement of the ventral prefrontal cortex is able to produce the binding of object characteristics that occurs in the working memory [van der Voort van der Kleij et al. 2003]. The memory model is developed using a feedforward network that includes regions for the cortex regions of V1, V2, V4, posterior infero-temporal cortex and the anterior infero-temporal cortex and a feedback mechanism that provides the identity of the object. In the visual working memory the activation in the top layer of the ventral prefrontal cortex is a copy of the activation in the posterior infero-temporal cortex. The lower level of the ventral prefrontal cortex is linked to the anterior infero-temporal cortex. These links are similar to those between the anterior inferior cortex and posterior infero-temporal cortex when performing a feedback role. These links are copied from the feedback network, with the depiction in this layer being the same as the representation of the posterior infero-temporal cortex.

O'Reilly and Frank (2006) and Hazy et al. (2006) also devised on computational model of working memory by modelling the prefrontal cortex and basal ganglia interaction and how a working memory model is able to perform sequential activities, in a similar manner to what occurs in the prefrontal cortex basal ganglia interaction to achieve action selection. This model depends on maintaining depictions in the prefrontal cortex which are updated/gated via the basal ganglia. When considering an architecture that models the functionality and interaction between the basal ganglia, prefrontal cortex and hippocampus, working memory is said to develop through these regions interactions including maintaining activity relevant information and fast learning of relations. Six main characteristics of working memory are identified by Hazy et al. (2006) that constrain this working memory model. These six are (i) working memory should represent and maintain new information as it happens; (ii) information that is required should be maintained despite the collection of other information; (iii) it must be possible to maintain separate representations in working memory; (iv) specific components of working memory are update while others remain unchanged; (v) there should be top-down control of other processes within the brain; (vi) it should be known when best to gate.

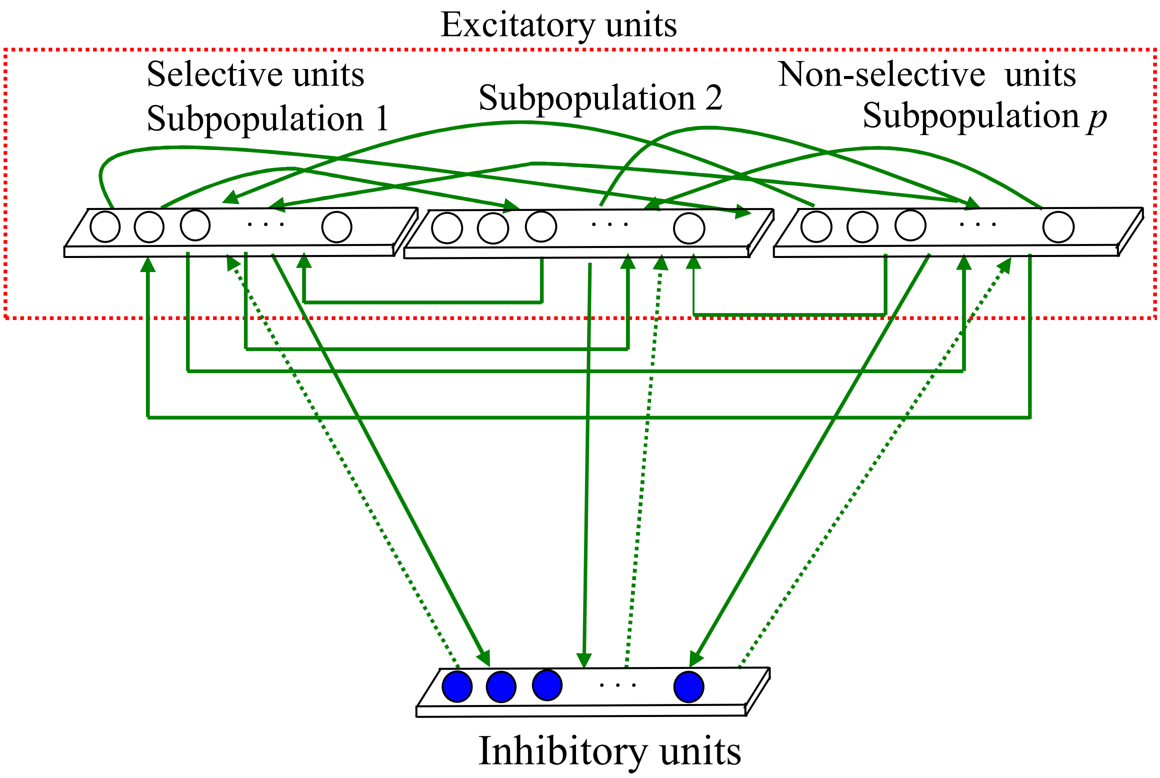


Figure 6.2.3 The working memory model after Omori and Horiguchi (2004). Solid lines are excitatory connections, dotted lines inhibitory connections .

A feature of the computational working memory model by Hazy et al. (2006) and O'Reilly and Frank (2006) is that the basal ganglia provides a gating approach to aid the maintenance of activation in the prefrontal cortex. Through a set of loops the basal ganglia and frontal cortex are linked, with neurons in the dorsal stratum causing the removal of inhibition on the frontal cortex to create gating like behaviour and so causing the prefrontal cortex to update the working memory depictions. The 'NoGo' neurons of dorsal striatum in the model through an indirect pathway countered this by inhibiting of the globus pallidus, external segment. The basal ganglia model determines if the information in working memory is relevant and so should be updated by use of a reinforcement learning actor-critic approach The basal ganglia is able to learn the relevant from the irrelevant working memory as a result of feedback coming from dopamine neurons. O'Reilly and Frank (2006) through their working memory model forecast that the information that is updated at different time intervals in the working memory activity is depicted in diverse locations in the prefrontal cortex.

The working memory model by O'Reilly and Frank (2006) was developed as a multitask model, with an input/output layer at the top of model, an hidden layer based on the posterior cortex, the prefrontal cortex in the centre and the bottom the basal ganglia that learns and gates the prefrontal cortex. The primary value learned approach uses two learning approaches which are a simple delta-rule or Rescorla-Wagner approach. Single or multiple inputs are introduced into strips in the Stimuli_In layer with the task instructions at the Task_Instruct and store/ignore/recall layers. Based on these inputs, plus context provided by prefrontal cortex input, the Hidden layer established the appropriate verbal or non-verbal output. The model has sections depicting the subcomponent of the prefrontal cortex and Matrix (striatal matrisomes) layers, and the elements of the substantia nigra pars compacta and SNrThal layers.

According to Hazy et al. 2006 in the multitask model through the store/ignore/recall it is possible to give explicit working memory update signals. The task instruction and store/ignore/recall input has hidden layers which allow them to create a depiction of the task requirements. These hidden layers with the perceptual input linked to a central hidden layer which is the posterior association cortex, which links with the output layers. The prefrontal cortex and the basal ganglia are linked to the high-level processing areas. To examine working memory O'Reilly and Frank (2006) modelled the 1-2-AX task, in which a subject after seeing a '1' must determine the consecutive letters 'A-X', but after seeing a '2' must determine the sequence 'B-Y'. The numbers '1', '2' are remembered for a greater time in an 'outer' loop. An 'inner' loop determines the required letter sequence within a short time. A third loop produced the motor output. It is noted that while the basal ganglia is only able to resolve these loops, the cortex is able to learn the contents within the loops. The model PFC stored them in 'stripes' with entries in a stripe being active using a winner-take-all approach.

7 Episodic memory

There are various types of memory. The two types of long-term that are usually distinguished are semantic and episodic memory. Semantic memory is information related to components of the environment such as objects, notions and concepts. Here, the location where this information is gained is not important [Hayes et al. 2004]. A fundamental element of the memory architecture for the ACORNS project is episodic memory. According to Rolls et al. (2002) episodic memory is the recollection of a particular event association with an occurrence. The event is something that occurs at a specific time and can have associated with it both spatial and non-spatial memories. Episodic memory replies on memorising events and situations based on who did what to whom and where this happened and can involve multiple objects, people etc. in a specific situation [Shastri 2001]. Episodic memory is associated with events like passing your driving test or moving to new house and so is an event that is typically associated with a spatial and temporal situation [Hayes et al. 2004]. From neuroscience studies the frontal and medial temporal areas are found to have a role in representing and retrieving episodic memory and the hippocampal system in representing and encoding the spatial-location memories [Hayes et al. 2004].

7.1 Biological basis of episodic memory

A study performed by Hayes et al. (2004) examines the regions of brains involved in episodic memory by showing subjects a video of the inside of various houses during which the objects and their location were highlighted. During fMRI scanning the subjects are shown pictures and asked questions such as 'which object did you observe', 'which scene did you observe', 'which object appeared first', 'which scene appeared first'. This study was able to show that when retrieving episodic memory activation is observed in the medial temporal, frontal, fusiform, and partietal areas. Certain regions are associated with object, spatial and temporal memories, while other concentrated on contextual memories and others location memories. For Hayes et al. (2004) the right parahippocampal gyrus is involved in the retrieval of location memories, location and temporal information is associated with the right dorsolateral prefrontal cortex and visual scene information related to the bilateral posterior parietal areas. Norman et al. (2006) notes that the research into episodic memory has focused on three brain regions: the hippocampus for retrieving information related to previous events, perirhinal cortex to determine familiar items; and the prefrontal cortex for memory targeting.

7.2 Computational models of episodic memory

A popular memory model of the form episodic memory takes is the complementary learning model [Norman et al. 2006, Norman and O'Reilly 2003]. The basis of this model is that the neocortex is used to create an internal memory of the environment that is adapted, learned incrementally and uses an overlapping representation for similar pattern. In contrast the hippocampus is able to learn quickly memory activations from the neocortex that can be retrieved quickly and as clearly distinct representations for all patterns. In the model according to Norman et al. (2006) episodic memory is achieved through the hippocampus retrieving specific details from memory and the neocortex assisting recognition by the use of a scalar familiarity level.

Norman and O'Reilly (2003) took the complementary learning model and developed it as a hippocampus and computational neocortex-episodic memory model. Learning in this model is achieved by using a Hebbian rule known as instar learning. Items to be memorised by the hippocampus and the neocortex are depicted using excitatory activations of units in the network, which are passed by excitatory synapses and controlled by inhibitory feedback. In Norman and O'Reilly (2003) the hippocampus model connects the input representation in the Entorhinal cortex with a group of units in the CA3 [see Figure 7.2.1]. The dentate gyrus is used to achieve feature extraction in the CA3. Through the use of recurrent connections it is possible to connection all units that are in CA3 that are used to represent an input from the Entorhinal cortex. The activation patterns produced by CA3 are feedback to the Entorhinal cortex via CA1. The network's connections and learning approaches allows it to retrieve full input patterns from the Entorhinal cortex using cues. According to Norman (2006) this model represented the two layers of the Entorhinal cortex as separate regions, one passes the input into the hippocampus and the other receives the output from the hippocampus. By combining these hippocampus and neocortex models Norman and O'Reilly (2003) are able to recreate some of the findings in neuroscience experiments related to list learning.

Although the hippocampus model is widely accepted, the neocortical model has been questioned [Norman et al. 2006]. One criticism of this model by Bogacz and Brown (2003) is it lack of memory storage capability. As a response to this problem Bogacz and Brown (2003) developed their own memory model of perirhinal familiarity determination. Although for the complementary learning model of Norman and O'Reilly (2003) relies on Hebbian feature extraction to achieve familiarity discrimination Bogacz and Brown model is based on an anti-Hebbian model. This model holds that neurons that perform familiarity determination use an anti-Hebbian approach that weakens the weights from active pre-synaptic to post-synaptic units and strengths this link for non-active units. The main gain from using such an anti-Hebbian approach is that memory network is able to detect novel features in a pattern.

According to Shastri (2001), episodic memory must fulfil certain criteria. For instance, the memory trace of episodic memory should ensure they represent the link between those involved in the event and what they do in this event. The memory trace should allow the differentiation between the memory and similar situations that are very close. Further, the episodic memory trace should allow the retrieval of particular elements of the event. According to Shastri (2001) if the memory trace is to fulfil these criteria there is a need for it to include functional elements that offer binding determination, binding-error determination, binding-error combination, association-instance-match-identifiers and binding-reinstators.

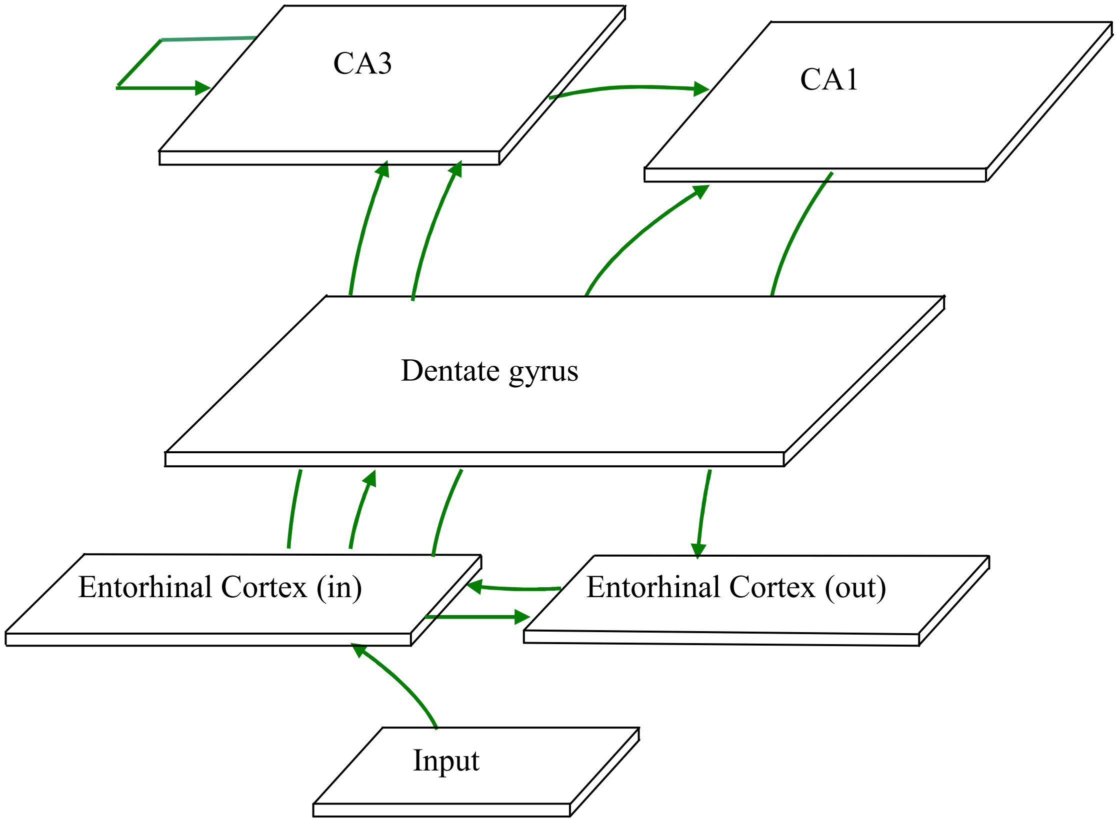


Figure 7.2.1 Representation hippocampus network of the complementary learning memory model. (From Norman and O'Reilly 2003)

Shastri (2001) developed a hippocampus based computational memory model that shows a transient activation pattern depicting an event can produce the appropriate functional units from long term potentiation. The model takes a stream of events which produce a representation of activations on the higher cortical regions, which project onto the Enorhinal cortex that produces a future representation of the activations. The activations in the Enorhinal cortex are passed in a loop through the dentate gyrus, CA3, CA2, CA1, subicular complex and return to the Enorhinal cortex to produce specific synapse alterations. These synapse alterations can take a transient activation pattern and create a memory trace that includes the required functional elements for episodic memory. The association between the functional elements making up a memory trace and the units of the hippocampus system are as follows: Neurons link the high-level cortical circuits depiction of entities, generic relational architectures and their function to the hippocampus system; binding-detector neurons in dentate gyrus; Binding-error-detector systems in CA3; binding-error-integrator neurons in CA2; relational-match-indicator systems in CA1 for identifying a match between a cue and the recalled event using the activity of the above neurons and systems; and binding-reinstator neurons in subicular complex.

Besides the biologically motivated models of episodic memory considered above, there are also more abstract models [Norman et al. 2006]. One such abstract memory model is known as MINERVA 2 and was developed by Hintzman (1988). In this model memory traces are placed in a store and are compared with the cue to determine the level of match between the cue and the trace. For MINERVA 2 a trace is depicted as a string of features, with 1 representing the feature is present, -1 it is missing and 0 the feature is not known. Similarity is determined using as the cue-trace dot product, divided by the number of characteristics included in the dot product:

$$s_i = \frac{\sum_{j=1}^{N} P_j T_{i.j}}{i_N}$$

s_i acts as the similarity value, P_j acts as the value of characteristic j in the cue, $T_{i,j}$ acted as the value of characteristic j in trace i, N is the number of characteristics, and N_i is the number of characteristics where the cue or trace are not zero.

In this model once the similarity score is determined for each trace they are cubed to develop an activation level A_i for each of the traces. Once this is done the activations are added together for all the traces to create echo intensity.

$$I = \sum_{i=1}^{m} A_i$$

The MINERVA 2 model is used by Maier and Moore (2005) on an automated speech recognition activity. The model is tested on the Peterson and Barney set which are readings for the first three formats and for the fundamental frequency from a group of vowel from 87 subjects. To accommodate the input frequency features the representation approach is extended to incorporate numerical values. The data from the set included in the feature vectors is the fundamental frequency, phoneme classes, gender and affiliation. In the recognition task Maier and Moore (2005) establish that the MINERVA 2 is able to outperform classification approaches such as Gaussian mixture models, K-nearest neighbour and support vector machines. According to Maier and Moore (2005) an interesting characteristic of MINERVA 2 is that feature vector include input and output information to incorporate context-associated characteristics to achieve episodic memory.

Another type of episodic memory models that has been developed is known as strategic memory search, which rather than having a cue requires the creation of your own cues. An example of this by Norman et al. (2006) is when a subject is not able to find their keys and so try and remember what they were doing earlier in the day. Various approaches to strategic memory search have been devised of which one is the free recall approach where individuals are required to memorise a word list and then recite it in any order. A model that is used to explain how people regain memory from a specific temporal context is the temporal context model. In this model the inner state of the person changes slightly over time. To achieve this according to Norman et al. (2006) Mensink and Raaijmakers (1988) use a binary context vector whose elements are updated at each time step, with the greater the probability of being updated the faster the change. In the memorising stage items are related to the context vector, while the reciting procedure is started by cuing the present values of the context vector which causes the regaining of item that are related to the contextual elements in the memorising stage. The temporal context models is different from contextual-drift models such as that of Mensink and Raaijmakers (1998) in that for the temporal context model's context drift occurs based on the characteristics of the item being memorised rather than occurring randomly.

8 ACORNS Memory architecture

As can be seen from Figure 8.1, which represents the preliminary memory model developed for the ACORNS projects we have incorporated the different features of memory models examined above. Central to the overall architecture is the memory-prediction framework of Hawkins and Blakeslee (2004), which will provide a form of hierarchical learning that matches neocortical region through an autoassociator network. The ACORNS memory architecture will include reinforcement memory learning based the responses from caregivers, which will be of the form of an actor-critic model. By using a reinforcement learning approach it will be possible to incorporate a predictive model to the language recognition and production. The ACORNS memory architecture will also include sensor/motor memory that offers the opportunity to ground spoken language using associator network including characteristics and evidence on how the brain performs. The memory model will combine a model of long-term memory in particular episodic memory with a representation of working memory. To some extent these two forms of memory will also take their inspiration from the biological system. For instance the episodic memory binds multimodal inputs to achieve an associator network that is able to take a memory fragment (a section of the input) and from this recreate the full memory. In the first instance, working memory is likely to take the form of recurrent feedback from previous states, as is typically found in simple recurrent networks. This will be built upon to offer a model of working memory that is closer to the suggestion of Baddeley (1992) by combining modules that act as the central executive, phonological loop, episodic buffer and visuospatial sketchpad. As seen in this study attention is fundamental if the system is not to be overloaded by too much stimuli, with it likely in our model that there will be a form of salience maps similar to the approach of Kayser et al. (2005) and form the basis of one of the roles played by the central executive in working memory model.

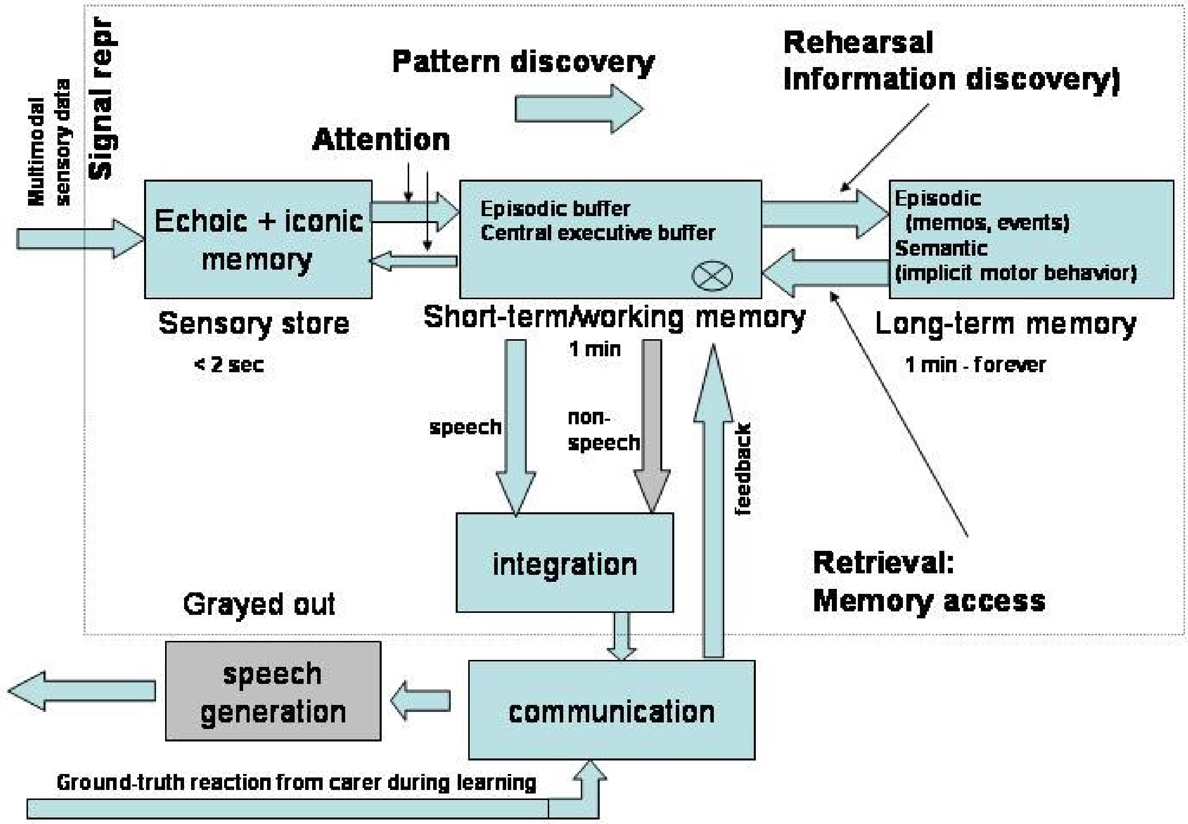


Figure 8.1 Preliminary memory model for ACORNS project

9 References

Amit, R., Matarić, M. (2002) Learning movement sequences from demonstration. Proceedings of the IEEE International Conference on Development and Learning (ICDL-2002), Cambridge, MA, USA, MIT Press, pp 203-208.

Arbib, M. (2005) From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. Behavioral and Brain Science, 28(2), pp. 105-124.

Arons, B. (1992) A review of the cocktail party effect. Journal of the American Voice I/O Society, 12, pp. 35-50.

Baddeley, A. (1992) Working memory. Science, 255(5044), pp. 556-559.

Baddeley, A. (2001) Is working memory still working? American Psychologist, 56, p.851-64.

Baddeley, A. D., Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory. 8, pp. 47–89.

Baddeley, A. (2003a) Working memory: Looking back and looking forward. Neuroscience, 4, pp. 829-839.

Baddley, A. (2003b) Working memory and language: An overview. Journal of Communication Disorders, 36, pp. 189-208.

Belpaeme, T., de Boer, B., de Vylder, B., Jansen, B. (2003) The role of population dynamics in imitation. Proceedings of the 2nd International Symposium on Imitation in Animals and Artifacts, Aberystwyth, University of Wales, pp. 171–175.

Billard, A. (1999) DRAMA, a connectionist architecture for on-line learning and control of autonomous robots: Experiments on learning of a synthetic proto-language with a doll robot. Industrial Robot Journal, 26(1), pp. 59-66.

Bodén, M. (2002) A guide to recurrent neural networks and backpropagation, in The DALLAS project. Report from the NUTEK-supported project AIS-8: Application of Data Analysis with Learning Systems, 1999-2001. Holst, A. (ed.), SICS Technical Report T2002:03, SICS, Kista, Sweden.

Bogacz, R., Brown, M. (2003) Comparison of computational models of familiarity discrimination in the perirhinal cortex. Hippocampus, 13, pp. 494-524.

Buccino, G., Vogt, S., Ritzi, A., Fink, G., Zilles, K., Freund, H.-J., Rizzolatti, G. (2004) Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. Neuron, 42, pp. 323-334.

Choi, S-B, Ban, S-W., Lee, M. (2004) Biological motivated visual attention system using bottom-up salience map and top-down inhibition. Neural Information Processing Letters and Review, 2(1), pp. 19-25.

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., Peterson, S. E. (1991). Selective and divided attention during visual discrimination of shape, color, and speed: Functional anatomy by positron emission tomography. J. Neurosci. 11, pp. 2383–2402.

Courtney, S., Petit, L., Haxby, J., Ungerleider, L. (1998) The role of prefrontal cortex in working memory: examining the contents of consciousness. Phil. Trans. R. Soc. Lond. B (1998) 353, pp. 1819-1828.

Cowan, N. (1999) An embedded-processes model of working memory. Models of working memory: Mechanisms of active maintenance and executive control. Miyake, A. and Priti S. (Eds). pp. 62-101.

Cowan, N. (2001) The Magical Number 4 in Short-term Memory: A Reconsideration of Mental Storage Capacity. Behavioral and Brain Sciences, 24, pp. 87-185.

Daw, N., Doya, K. (2006) The computational neurobiology of learning and reward, Current Opinion in Neurobiology, 16, pp. 199–204.

Dayan, P. and Balleine, B. (2002) Reward, Motivation, and Reinforcement Learning. Neuron (36), pp. 285–298.

Dayan, P., Hinton, G. (1996) Varieties of Helmholtz machine. Neural Network, 9(8), pp. 1385-1403.

Dayan, P., Hinton, G. E., Neal, R., Zemel, R. S. (1995) The Helmholtz machine. Neural Computation, 7, pp. 1022-1037.

Dayan, P., Hinton, G. E., Neal, R., Zemel, R. S. (1995) The Helmholtz machine. Neural Computation, 7, pp. 1022-1037.

Demiris, Y., Hayes, G. (2002) Imitation as a dual-route process featuring prediction and learning components: A biologically-plausible computational model. Imitation in animals and artifacts, Dautenhaln, K., Nehaniv, C., Eds., Cambridge, MA, USA, MIT Press, pp. 327-361.

Demiris, Y., Hayes, G. (2002) Imitation as a dual-route process featuring prediction and learning components: A biologically-plausible computational model. Imitation in animals and artifacts, Dautenhaln, K., Nehaniv, C., Eds., Cambridge, MA, USA, MIT Press, pp. 327-361.

Dillmann, R. (2003) Teaching and learning of robot tasks via observation of human performance. Proceedings of IROS-03 Workshop on Programming by Demonstration, pp. 4-9.

Egner, T., Hirsch, J. (2005) Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. Nature Neuroscience, 8, pp. 1784-1790.

Elman, J. (1990) Finding structure in time. Cognitive Science, 14, pp. 179-211.

Gallese, V. (2005) The Intentional attunement hypothesis. The mirror neuron system and its role in interpersonal relations. Biomimetic Multimodal Learning in a Robotic Systems, Wermter, S., Palm, G., Elshaw, M. (Eds), Heidelberg, Germany, Springer-Verlag.

Gallese, V., Escola, L., Intskiveli, I., Umilta, M., Rochat, M., Rizzolatti, G. (2004) Goal-relatedness in area F5 of the macaque monkey during tool. MirrorBot Technical Report 17.

Gallese, V., Goldman, A. (1998) Mirror neurons and the simulation theory of mind-reading. Trends in Cognitive Science, 12(2), pp. 493-501.

Garalevicius, S. (2007) Memory prediction framework for pattern recognition: Performance and suitability of the Bayesian Model of visual cortex. 20th International FLAIRS Conference. Key West, Florida.

Garfield S., Wermter S. (2006) Call Classification using Recurrent Neural Networks, Support Vector Machines and Finite State Automata. Knowledge and Information Systems: An International Journal, 9(2), pp. 131-156.

Garfield S., Wermter S. Recurrent Neural Learning for Helpdesk Call Routing. Proceedings of the International Conference on Artificial Neural Networks, Madrid, Spain, pp. 296-301, August 2002.

Gerlach, G., Aaside, C., Humphreys, G., Gade, A., Paulson, O., Law, I. (2002) Brain activity related to integrative processes in visual object recognition: Bottom-up integration and the modulatory influence of stored knowledge. Neuropsychologia, 40(8), pp. 1254-1267.

George, D., Hawkins, J. (2005) A Hierarchical Bayesian Model of Invariant Pattern Recognition in the Visual Cortex. In Proceedings of the International Joint Conference on Neural Networks, Montreal, Canada.

Grabowski, T., Damasio, H., Damasio, A. (1998) Premotor and prefrontal correlates of category-related lexical retrieval. Neuroimages, 7, pp. 232-243.

Gurney, K. N., Prescott, T. J., Redgrave, P. (2001a). A Computational Model of Action Selection in the Basal Ganglia. I. A new functional anatomy. Biological Cybernetics. 84, pp. 401-410.

Gurney, K. N., Prescott, T. J., Redgrave, P. (2001b). A Computational Model of Action Selection in the Basal Ganglia. II. Analysis and simulation of behavior. Biological Cybernetics. 84, pp. 411-423.

Harnad, S. (1990) The symbol grounding problem. Physica D, 42, pp. 335-346.

Harnad, S. (2003) The symbol grounding problem. Encyclopedia of Cognitive Science, London, Nature Publishing Group, Macmillan.

Hauk, O., Pulvermüller, F. (2004) Neurophysiological distinction of action words in the fronto-central cortex. Human Brain Mapping, 21, pp. 191–201.

Hauk, O., Pulvermüller, F. (2004) Neurophysiological distinction of action words in the fronto-central cortex. Human Brain Mapping, 21, pp. 191–201.

Hauk, O., Pulvermüller, F. (2004) Neurophysiological distinction of action words in the fronto-central cortex. Human Brain Mapping, 21, pp. 191–201.

Hawkins, J., Blakeslee, S. (2004) On Intelligence. New York, NY, Times Books.

Hazy, T., Frank, M., O'Reilly, R. (2006) Banishing the homunculus: Making working memory work. Neuroscience 139, pp.105–118.

Hinton, G.E., Dayan, P., Frey, B.J., Neal, R. (1995) The wake-sleep algorithm for unsupervised neural networks. Science, 268, pp 1158-1161.

Hintzman, D. (1986) Schema-Abstraction in a multiple-trace memory model. Psychology review, 93, pp. 411-427.

Houk, J. C., Adams, J. L., Barto, A. G. (1995). A model of how the basal ganglia generate and use reward signals that predict reinforcement. In Houk, J., Davis, J., Beiser, D. (Eds.), Models of information processing in the basal ganglia, pp. 249-270, Cambridge: MIT Press.

Hyes, S., Ryan, L., Schnyer, D. and Nadel, L. (2004) An fMRI study of episodic memory: Retrieval of objects, spatial, and temporal information. Behavioral Neuroscience, 118(5), pp. 885-896.

Iwasaki, M. Hashiyama, T. Okuma, S. (1999) Extraction of features and attention areas in hierarchical neural networks. In proceedings of Systems, Man, and Cybernetics.

Joel, D., Niv, Y., Ruppin, E. (2002) Actor-critic models of the basal ganglia: new anatomical and computational perspectives. Neural Networks, 15(4), pp. 535-547.

Kayser, C., Petkov, C., Lippert, M. and Logothetis, N. (2005) Mechanisms for Allocating Auditory Attention: An Auditory Saliency Map, Current Biology, Vol. 15, 1943–1947, November 8, 2005.

Khamassi, M. , Lachèze, L., Girard, B., Berthoz, A., Guillot, A. (2005). Actor-critic models of reinforcement learning in the basal ganglia: From natural to artificial rats. Adaptive Behavior, Special Issue Towards Artificial Rodents, 13(2):131-148.

Kiefer, M. (2001) Perceptual and semantic sources of category-specific effects: Event-related potentials during picture and word categorization. Memory and Cognition, 29, pp. 100-116.

Kohler, E., Keysers, C., Umilta, M., Fogassi, L., Gallese, V., Rizzolatti, G. (2002) Hearing sounds, understanding actions: Action representation in mirror neurons. Science, 297, pp. 846-848.

Maier, V. and Moore, R. (2005) An investigation into a simulation of episodic memory for automatic speech recognition. In Proceedings of ISCA INTERSPEECH.

Mavridis, N. and Roy, D. (2006). Grounded Situation Models for Robots: Where Words and Percepts Meet. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).

Meeter, M., Myers, C. and Gluck, M. (2005) Integrating incremental learning in episodic memory models of the hippocampal region. Psychological Review, 112, pp. 560-585.

Mensink, G. and Raaijmakers, J. (1998) A model for interference and forgetting. Psychological review, 95, pp. 434-455.

Mizuno, A. (2005) Process Model for Simultaneous Interpreting and Working Memory. Processus et cheminements en traduction et interprétation, 50(2).

Moore, R. K. (2007) Spoken language processing: Piecing together the puzzle. Speech communication, 49, pp. 418-435.

Mountcastle, V. (1978) An organizing principle for cerebral function the unit model and the distribution system. In Edelman, G. and Mountcastle, V. (Eds) The Mindful brain. MIT press.

Nieuwenhuis, S. and Yeung, N. (2005) Neural mechanisms of attention and control: Losing our inhibitions. Nature Neuroscience, 8(12), 1631-1633.

Norman, K. and O'Reilly, R. (2003) Modeling hippocampal neocortical contributions to recognition memory: A complementary-learning-systems approach. Psychological Review, 104, pp. 611-646.

Norman. K., Detre, G. and Polyn, S. (2006) Computational models of episodic memory. The Cambridge Handbook of Computational Cognitive Modelling. Sun, R. (Eds).

Omari, T. and Horiguchi, T. (2004) Dynamic properties of neural network model for working memory with Hodgkin-Huxley neurons. Physica A, 334, p. 600-614.

O'Reilly, R., Frank, M. (2006) Making Working Memory Work: A Computational Model of Learning in the Prefrontal Cortex and Basal Ganglia. Neural Computation, 18(2), pp. 283-328.

Papagno, C., Valentine, T. and Baddeley, A. (1992) Phonological short-term memory and foreign language vocabulary learning. J. Mem. Lang. 30, pp. 331-347.

Pardo, J. V., Fox, P. T., and Raichle, M. E. 1991. Localization of a human system for sustained attention by positron emission tomography. Nature 349, pp. 61–64.

Perani, D., Cappa, S., Schnur, T., Tettamanti, M., Collina, S., Rosa, M., Fazio, F. (1999) The neural correlates of verbs and noun processing a PET study. Brain, 122, pp. 2337-2344.

Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. Annu. Rev. Neurosci. 13, pp. 25–42.

Posner, M. I., Presti, D. (1987) Selective attention and cognitive control. Trends Neurosci. 10, pp. 12–17.

Priti, S. and Miyake, A. (1999) Models of Working Memory: An Introduction. Models of working memory: Mechanisms of active maintenance and executive control. Miyake, A. and Priti S. (Eds).

Prommer, T., Holzapfel, H. and Waibel, A. (2006) Rapid simulation-driven reinforcement learning of multimodal dialog strategies in human-robot interaction. In proceedings of Interspeech 2006.

Pugh, K., Shaywitz, B., Shaywitz, S., Fulbright, R., Byrd, D., Skudlarski, P., Shankweiler, D., Katz, L. Constable, R., Fletcher, J., Lacadie, C., Marchione, K., Gore, J. (1996) Auditory Selective Attention: An fMRI Investigation, Neuroimage, 4, pp. 159–173.

Pulvermüller, F. (1999) Words in the brain's language. Behavioral and Brain Sciences, 22(2), pp. 253-336.

Pulvermüller, F. (2001) Brain reflections of words and their meaning. Trends in Cognitive Neuroscience, 12(5), pp. 517-524.

Pulvermüller, F. (2002) A brain perspective on language mechanisms: from discrete neuronal ensembles to serial order. Progress in Neurobiology, 67, pp. 85-111.

Pulvermüller, F. (2003) The neuroscience of language: On brain circuits of words and language. Cambridge, UK, Cambridge Press.

Pulvermüller, F., Assadollahi, R., Elbert, T. (2001) Neuromagnetic evidence for early semantic access in word recognition. European Journal of Neuroscience, 13, pp. 201-205.

Pulvermüller, F., Birbaumer, N., Lutzenberger, W., Mohr, B. (1997) High-frequency brain activity: Its possible role in attention, perception and language processing, Prog. Neurobiol, 52, pp. 427-445.

Pulvermüller, F., Häre, M., Hummel, F. (2000) Neurophysiological distinction of verb categories. NeuroReport, 11, pp. 2789-2793.

Pulvermüller, F., Mohr, B., Schleichert, H. (1999) Semantic or lexico-syntactic factors: What determines word class specific activity in the human brain? Neuroscience Letters, 275(81-84), pp. 81-84.

Pulvermüller, F., Moscoso-del-Prado-Martin, F., Hauk O. (2004) Category specificity in the processing of colour and form related words. Technical Report 13.

Rizzolatti, G. (2002) Hearing sounds, understanding actions: Action representation in mirror neurons. Science, 297, pp. 846-848.

Rizzolatti, G., Arbib, M. (1998) Language within our grasp. Trends in Neuroscience, 21(5), pp. 188-194.

Rizzolatti, G., Fogassi, L., Gallese, V. (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Review, 2, pp. 661-670.

Rizzolatti, G., Luppino, G. (2001) The cortical motor system. Neuron, 18(2), pp. 889-901.

Rolls, E., Stringer, S., Trappenberg, T. (2002) A unified model of spatial and episodic memory. Proc. R. Soc. Lond. B, 269, pp. 1087-1093.

Roy, D. (2003) Grounded spoken language acquisition: Experiments in word learning. IEEE Transactions on Multimedia, 5(2), pp. 197- 209.

Roy, D. (2004) Grounding language in the world: Signs, schemas, and meaning. Artificial Intelligence.

Sauser, E. and Billard, A. (2005) View sensitive cells as a neural basis for the representation of others in a self-centered frame of reference. Proceedings of the Third International Symposium on Imitation in Animals and Artifacts, Hatfield, UK.

Schaal, S., Ijspeert, A., Billard, A. (2003) Computational approaches to motor learning by imitation. Transaction of the Royal Society of London: Serial B, Biological Sciences, 358, pp. 537-547.

Shastri, L. (2001) A computational model of episodic memory formation in the hippocampal system. Neurocomputing, 38-40, pp. 889-897.

Shtyrov, Y., Hauk, O., Pulvermüller, F. (2004) Distributed neuronal networks for encoding category-specific semantic information: the mismatch negativity to action words. European Journal of Neuroscience, 19, pp. 1-10.

Singh, S., Litman, D., Kearns, M., Walker, W. (2002) Optimizing Dialogue Management with Reinforcement Learning: Experiments with the NJFun System by Satinder Singh, Diane Litman, Michael Kearns and Marilyn Walker. Journal of Artificial Intelligence Research (JAIR), 16, pp. 105-133, 2002.

Sun, R. (2000) Symbol grounding: A new look at an old idea. Philosophical Psychology, 13(2), pp.149-172.

Suri, R. E., Schultz, W. (1998). Learning of sequential movements by neural network model with dopamine-like reinforcement signal. Exp. Brain Res., 121, 350-354.

Suri, R. E., Schultz, W. (1999). A neural network model with dopamine-like reinforcement signals that learns a spatial delayed response task. Neuroscience, 91, 871-890.

Suri, R. E., Bargas, J., Arbib, M. A. (2001). Modeling functions of striatal dopamine modulation in learning and planning. Neuroscience, 103, 65-85.

Tani, J., Ito, M., Sugita, Y. (2004) Self-organization of distributedly represented multiple behavior schemata in a mirror system: Reviews of robot experiments using RNNPB. Neural Networks, 17, pp.1273-1289.

Umilta, M., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., Rizzolatti, G. (2001) I know what you are doing: A neurophysical study. Neuron, 31, pp. 155-165.

van der Voote van der Kleij, G., de Kamps, M. and van der Velde, F. (2003) A neural model of binding and capacity in visual working memory. Proceedings of ICANN/ICONIP, pp. 771-778.

Vogt, P. (2006) Language evolution and robotics: Issues in symbol grounding and language acquisition. Loula, A., Gudwin, R., Queiroz, J. (Eds), Artificial Cognition Systems. Idea Group.

Weber C., Wermter S., Zochios A. Robot Docking with Neural Vision and Reinforcement. *Knowledge Based Systems*, Vol. 12, No. 2-4, pp. 165-72, 2004.

Wermter, S., Elshaw, M., Farrand, S. (2003) A Modular Approach to Self-organisation of Robot Control Based on Language Instruction. Connection Science, 15(2-3) pp. 73-94.

Wermter S., Weber C., Elshaw M., Gallese, V., Pulvermüller F. Grounding Neural Robot Language in Action. In Wermter S., Palm G., Elshaw M. Biomimetic Neural Learning for Intelligent Robots, pp. 162-181, 2005.

Winter, J., Schreiner, C. (2005) Inferior Colliculus. Springer.

Wrigley, S., Brown, G. (2004) A computational model of auditory selective attention. IEEE Transactions on Neural Networks, 15(5), pp.1151-1163.

Wrigley, S., Brown, G. (2002) A neural oscillator model of auditory selective attention. In Advances in Neural Information Processing Systems 14, T.G. Dietterich, S. Becker and Z. Ghahramani, (Eds) MIT Press

Wu, W., Kwasny S., Kalman, B., Engebretson, A. (1993) Identifying language from raw speech - An application of recurrent neural networks. In Proceedings 5th Midwest Artificial Intelligence and Cognitive Science Conference.

Zatorre, R., Mondor, T., Evans, A. (1999) Auditory attention to space and frequency activates similar cerebral systems. NeuroImage 10, pp. 544–554