2006 International Joint Conference on Neural Networks
Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada
July 16-21, 2006

Language Acquisition and Symbol Grounding Transfer with Neural Networks and Cognitive Robots

Angelo Cangelosi, Emmanouil Hourdakis, and Vadim Tikhanoff

Abstract— **Neural networks have been proposed as an ideal cognitive modeling methodology to deal with the symbol grounding problem. More recently, such neural network approaches have been incorporated in studies based on cognitive agents and robots. In this paper we present a new model of symbol grounding transfer in cognitive robots. Language learning simulations demonstrate that robots are able to acquire new action concepts via linguistic instructions. This is achieved by autonomously transferring the grounding from directly grounded action names to new higher-order composite actions. The robot's neural network controller permits such a grounding transfer. The implications for such a modeling approach in cognitive science and autonomous robotics are discussed.**

I. Introduction

LINGUISTIC communication is not an isolated capability of individuals, but rather it is intrinsically linked with other cognitive and sensorimotor capabilities [1,2,3,4]. We come to realize and comprehend language through a series of semantic interpretations of symbols and meanings within our world. Furthermore, these symbols do not exist as arbitrary representations of some notion, but are intrinsically connected to behavioral or cognitive abilities, based on the properties of the reference system they belong to. This task of connecting the arbitrary symbols used in internal reasoning with external physical stimuli is known as the "Symbol Grounding problem" [5].

In the next sections we will discuss the theoretical implications for the symbol grounding problem and the use of cognitive robotics to deal with it. We then propose a new simulation model for the symbol grounding transfer in agents that acquire a simple lexicon via imitation and linguistic instructions. This model is based on the combination of neural network and cognitive robotics methodologies.

This work was in part sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA8655-05-1-3060. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon.

A. Cangelosi is with the Adaptive Behaviour and Cognition research group, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK (phone: 44-1752-232559; fax: 44-1752-232540; e-mail: acangelosi@plymouth.ac.uk).

E. Hourdakis was with the Adaptive Behaviour and Cognition research group, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK (e-mail: ehourdakis@yahoo.com).

V. Tikhanoff is with the Adaptive Behaviour and Cognition research group, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK (e-mail: vtikhanoff@plymouth.ac.uk).

A. The symbol grounding problem

Language reference systems consist of objects and their associations. Peirce [6] defined three types of referential systems, namely icon, index and symbol. Icons regard representations that are directly connected to the objects they refer to according to their "conventional similarity" and stimulus generalization [7]. Indexical associations are found in animal communication systems, and are based on learned spatio-temporal associations between the referent and its index. Symbols, in addition to being associated to external referents, are characterized by their logical and combinatorial relationships with other symbols [5,7,8]. These symbols are not innate in humans, but are acquired and grounded through a series of stimuli and cognitive interaction with the world, embodied as words within a human communication system. Barsalou [9] further states that the brain implements these basic symbolic operations by predicating conceptual properties of individuals and categories.

One fundamental issue in symbolic cognitive systems is how grounded symbols acquire their meaning. Symbolic accounts of cognition (e.g. [10]) suggest that symbols are defined by their relation with other symbols, and so on. However, such a mechanism is subject to the infinite regression problem. To solve this, a model needs to ensure that its reference mechanism will eventually lead to one or more objects, or cognitive representations (e.g. categories), directly connected (grounded on) the world. This is the embodiment approach to cognition [3].

Harnad [5] has identified two mechanisms for grounding symbols. The first, called "sensorimotor toil", implies that the agent acquires new grounded symbols through direct sensorimotor interaction with the environment, under the guidance of corrective feedback [6]. Among the most important drawbacks of this method is the amount of training required for an agent to complete learning. Furthermore, categories acquired this way lack of any symbolic representations, thus restricting themselves only to iconic and indexical ones. In contrast to sensorimotor toil, "symbolic theft" [11] allows the acquisition of new concepts, and their symbolic representations, purely through the hearsay of propositional combinations of previously grounded symbols. In symbolic theft, the symbol grounding transfer mechanism permits the transfer of grounding from directly grounded symbols to new words. For example, if an individual has learned the meaning of the symbols "horse"

0-7803-9490-9/06/$20.00/

and "stripes" through direct experience with their referents (horses and striped patterns), she can also indirectly ground the meaning of the symbol "zebra" from a propositional definition such as "a zebra looks like a horse with stripes". This is possible through the transfer of the grounding from the basic symbols to the newly acquired word.

B. Neural network and cognitive robotic approaches to the symbol grounding problem

Neural networks have been proposed as an ideal cognitive modeling methodology to deal with the symbol grounding problem [5]. For example, feedforward multi-layer perceptron (MLP) models of the acquisition of lexicon (e.g. [12,13]) permit a good implementation of the process of grounding output symbolic representations in the (analogical) input representation of external stimuli. This is similar to the "sensorimotor toil" approach to grounding.

The same feedforward models can be extended to simulate the process of grounding transfer. For example, Cangelosi et al. [14,15] first trained an MLP to acquire output linguistic representation of visual input (e.g. geometric shapes or animal shapes). The same network is also able to receive symbols in input and reproduce in output a categorical representation of the object the symbols refer to. That is, the same network can perform linguistic production and comprehension. The MLP is then trained to acquire new symbols through the input of linguistic combinations of previously acquired basic words (e.g. "zebra = horse & stripes" in [4]). A grounding test stage demonstrates that these new concepts also acquires their grounding from that of the basic words "horse" and "stripes". This simulation implements the "symbolic theft" approach to grounding.

More recently, such neural network approaches have been incorporated in studies based on cognitive agents and robots. Cognitive robotics refers to the field of robotics that aims at builds autonomous cognitive systems capable of performing cognitive tasks such as perception, categorization, language and sensorimotor problems. Cognitive robotic approaches include epigenetic robotics [13], autonomous mental development systems [16], and evolutionary robotics [17,18]. These aim at modeling the developmental and evolutionary processes in the emergence of cognitive abilities in natural and artificial cognitive systems. Cognitive agent approaches refer to simulation models based on multi-agent systems and artificial life studies [15,19,20,21,22].

One cognitive agent model based on artificial life simulations has directly deal with the symbol grounding problem [11]. It uses a population of foraging agents that have to learn words to describe to each linguistically the different types of foods ("mushrooms"). The sensorimotor (i.e. navigation) and linguistic behavior of the agents is controlled by an MLP, as in [14], to perform symbol grounding and grounding transfer. This model is used to demonstrate the language origin hypothesis on the adaptive advantage of the "symbolic theft" mechanism versus the "sensorimotor toil" method for category learning. When two populations of foragers are put in direct competition with each other (toilers vs. thieves), symbolic theft always prevails.

In this paper we present a further application of the cognitive robotics approach to symbol grounding. The new model is an extension of a previous cognitive robotic research on language acquisition and symbol grounding transfer via direct instructions by humans [21]. In the preliminary model, robots are only able to perform linguistic comprehension tasks. That is, they receive in input only the names of action and produce in output the corresponding behavior. This is implemented through a feedforward multi-layer perceptron with linguistic-only input units and motor-only output units. In this new study we extend the model by implementing both linguistic comprehension and production. Robots are able to respond to a linguistic instruction and at the same time to produce a simple linguistic description of actions. The learning protocol follows the symbol grounding and transfer protocol originally proposed by Cangelosi et al. [14]. This consists of a three stage process, in which agents first learn to perform a basic action on an object, then associate the action and object with linguistic signals (name of actions and of objects), and later uses these signals to form higher order propositions describing new actions.

II. SIMULATION MODEL

A. The robot

Our system consists of two simulated agents (teacher and learner) embedded within a virtual simulated environment (Fig. 1). Each robot consists of two 3-segment arms attached to a torso. This is further connected to a base with four wheels. The robot has a total of 10 Degrees of Freedom (DOFs): shoulder joint (2: one for left and one for right arm), upperarm joint (2), elbow joint (2), wheels (4). In the current simulation, only 6 DOFs will be used, since the wheels are not used. Through the two arms the robot can interact with the environment and manipulate objects put in front of it. Three objects were used in the current simulation: a cube, a horizontal plane and a vertical bar. The agent can receive in the input retina different views (perspectives) of each object. The agent has to learn six basic actions: lower right shoulder, lower left shoulder, close right upperarm, close left upperarm, close right elbow, close left elbow. They will also learn the name of such basic actions: "LOWER_RIGHT_SHOULDER", "LOWER_LEFT_SHOULDER", "CLOSE_RIGHT_UPPERARM", "CLOSE_LEFT_UPPERARM", "CLOSE_RIGHT_ELBOW", "CLOSE_LEFT_ELBOW" Each action will be associated with some of the above objects that are put in front of the agent. The close left and close right shoulder actions are associated with different views of the cube.

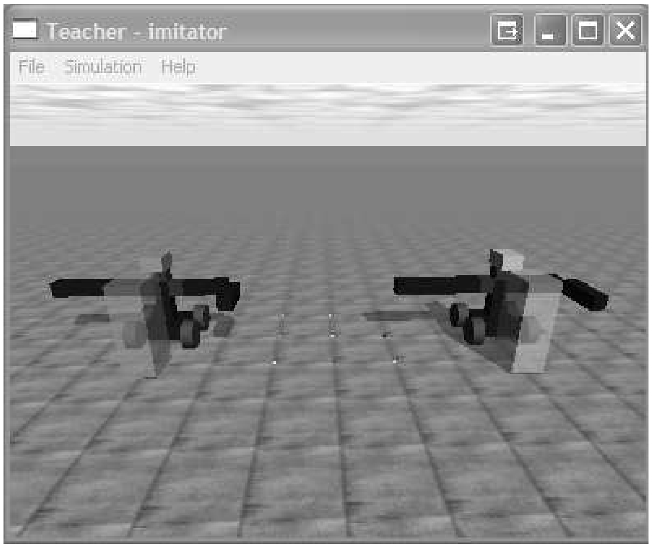

Fig. 1: Simulation setup with the two robots. The teacher robot is on the left and the learner on the right. The agents are performing the close left elbow action.

This system is implemented using ODE (Open Dynamics Engine, www.ode.org), an open source, high performance library for simulating rigid body dynamics. ODE is useful for simulating vehicles, objects in virtual reality environments and virtual creatures, and is being increasingly used for simulation studies on autonomous cognitive systems.

The first agent, the teacher, is pre-programmed to perform and demonstrates a variety of basic actions, each associated with a linguistic signal. These are demonstrated to the second robot, the learner, which attempts to reproduce the action by mimicking it. Firstly the agent acquires basic actions by observing the teacher. It then learns the basic action names (direct grounding). Subsequently, it autonomously uses the linguistic symbols that were grounded in the previous learning stage to acquire new higher-order actions (symbol grounding transfer).

B. Neural network controller

The imitator robot is endowed with a neural network that governs all of its perceptual, cognitive and motor abilities (Fig. 2). The neural network consists of a three layer feedforward MLP. There are three modalities integrated in input to the network, namely vision, motors and language. The vision consists of 25 input units for a 25 grid retina. This provides a simplified view of the object put in front of the agent. The motor input consists of 6 units for the proprioceptive information of the previous activation of the motor actuators. The linguistic input layer consists of 12 localist nodes, 6 referring to the names of the basic actions, 3 reserved for the higher order action names, and three for the names of the object.

The output layer consists of 6 motor units activating the 6 joint actuators. The node activation encodes the force that is being applied on the joint. Each action consists of a sequence of 10 steps of motor activations. The output value of each motor neuron is rescaled to the interval ±1. This corresponds to the force applied to each joint.

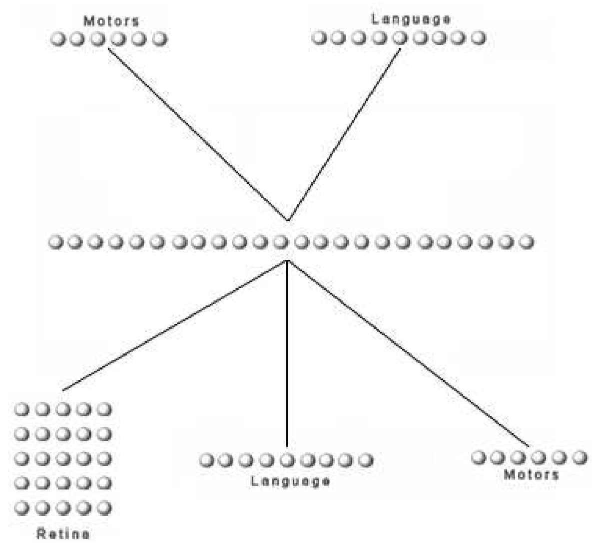

Fig 2. Architecture of the learner robot's neural network controller.

C. Training procedure

We attain the grounding transfer, using a 3 stage training process: (1) BA Basic Action learning, (2) EL Entry-Level naming and (3) HL Higher-Order learning.

During the Basic Action learning stage, the agent learns to execute all six basic actions in association with the view of the different objects. No linguistic elements are used at this stage. The imitation algorithm is used to adjust the weights contributing to the activation of the motor units. The algorithm is defined by the following functions:

$$f(t+1) = f(t) + g(x(t), y(t))$$

$$g(x(t), y(t)) = \alpha\left(\frac{2}{1+\exp(-2\beta(x(t)-y(t))}\right) - 1$$

with $\alpha = scale\ \beta = gain$.

The first function computes an estimation of the necessary force *f(t + 1)* to apply to each motorized joint in the next time-step, so that it approximates the posture currently exhibited by the demonstrator. It takes as input the joint angles *x(t)* of the demonstrator agent and the joint angles *y(t)* and motor forces *f(t)* of the imitator agent for all joints in the current time-step. Experimental evidence has demonstrated that joint angles are used for postural control in imitation [22]. The scale α and gain β are constant values, set to 0.5 in the present simulation. The scale parameter α is similar to the learning rate in the error backpropagation algorithm, where higher values produce bigger weight changes and faster learning. The gain parameter β changes the hyperbolic function (lower values correspond to flatter sigmoids).

The second learning stage, Entry Level naming (EL), was concerned with associating the previously acquired behaviors to linguistic signals. It features three sequential activation cycles. The first EL cycle, Linguistic Production, trains the learner how to name the 6 basic actions. Motor (proprioceptive) and visual (object view) information are given in input to the network. The agents learn to correctly activate the output linguistic nodes corresponding to the basic action names. This is based on a standard backpropagation algorithm. This linguistic production cycle implements the process of basic symbol grounding, by which the names (symbols) that the agent is learning are directly grounded on its own perceptual and sensorimotor experience. In the second EL cycle, Linguistic Comprehension, learner agents are taught to correctly respond to a linguistic signal consisting of the name of the action, without having the ability to perceive the object associated to the action. To accomplish this, the retinal units in the network were set to 0, whilst we activate the input units corresponding to the action name. In the final EL cycle, Imitation, both motor and linguistic inputs were activated in input, and the network learns to reproduce the action in output and activate the corresponding action name unit. This third cycle is necessary to permit the linking of the production and the comprehension tasks in the hidden units activation pattern (see [14]).

The final training stage, Higher-Level (HL) learning, allows the learner agents to autonomously acquire higher-order actions without the need of a demonstration from the teacher. This is achieved only through a linguistic instruction strategy and a "mental simulation" strategy similar to Barsalou's perceptual symbol system hypothesis [9]. The teacher only has to provide new linguistic instructions consisting of the names of two basic actions and the name of a new higher-order action.

Agents learn three higher-order actions with the following linguistic instructions:

"LOWER_RIGHT_SHOULDER+LOWER_LEFT_SHOULDER=PLACE",
"CLOSE_RIGHT_UPPERARM+CLOSE_LEFT_UPPERARM=HUG",
"CLOSE_RIGHT_ELBOW+CLOSE_LEFT_ELBOW=GRAB".

Once the teacher (or a human instructor) provides a higher-order instruction, the learner goes through four HL learning cycles. First it activates only the input unit of the first basic action name to produce and store ("memorize") the corresponding sequence of 10 motor activation steps. Secondly, it activates in input the linguistic units for the first basic action name and the new higher-order action. The resulting 10 motor activations are compared with the previously stored values to calculate the error and apply the backpropagation weight corrections. The next two cycles are the same as the first two, with the difference of activating the second basic action name unit.

Two different simulations of the Higher-Order learning stage were carried out. In the first version, HL1, only the name of the new action is given in input in the 2nd and 4th cycles described above. In the second alternative, HL2, the agent activates in input simultaneously the new higher-order name and the basic-order name for the same cycles 2 and 4.

The Higher-Order stage permits the implementation of a purely autonomous way to acquire new actions through the linguistic combination of previously-learned basic action names. The role of the teacher in this stage is only that of providing a linguistic instruction, without the need to give a perceptual demonstration of the new action. The motor imitation learning, such as in the Basic Action training stage, is a slow process based on continuous supervision, trial-and-error feedback-based learning. The acquisition of a new concept through linguistic instruction is, instead, a quicker learning mechanism because it requires the contribution of fewer units (the localist linguistic units) and corresponding weights. Moreover, in a related symbol grounding model on language (symbolic) vs. error-and-trial (sensorimotor toil) learning of categories, the linguistic procedure consistently outperforms the other learning method [11].

To establish if the agent has actually learned the new high-order actions and transferred the grounding from basic action names to higher order names, a test phase is performed. This grounding transfer test aims at evaluating the aptitude of the imitator agent to perform a new composite action with any of the objects previously associated, in the absence of the linguistic descriptions of the basic actions. Thus the agent is requested to respond solely on the signal of the composite action (e.g. Grab) and selectively to the different view of the objects. In addition, while the imitator was taught only the motion of the dissected action for each composite behavior, the test evaluated the performance of the higher-order composite action. This was a behavior never seen before by the robot. The stage was comprised of two basic trials per behavior, using the different view of the objects. All inputs were propagated through the network with no training occurring. In addition, two testing phases were run, one for each alternative trained during the previous stage HL1 and HL2.

III. RESULTS

We replicated the simulation experiment as above with five agents. Each agent had a different set of random weights initialized in the range ±1. The three learning stage, Basic Action, Entry-Level and Higher-Level learning, respectively lasted for 1000, 3000, 1500 epochs. This was the approximate minimum number of epochs necessary to reach a good learning performance. The parameters of the backpropagation algorithm were set as follow: BA stage, momentum 0.6 and learning rate 0.2; EL stage, momentum 0.6 and learning rate 0.3; HL stage, momentum 0.8 and learning rate 0.2 . The weights were updated at the end of every action.

Overall, results indicate that all agents were able to successfully learn the 6 basic actions and 3 higher-order behaviors. The three charts in Fig. 3 below provide an example of the learning curves for the BA/EL/HL1 stages in the agent with the lowest final error. The variability between the five replications was very low, around ±0.01 error.

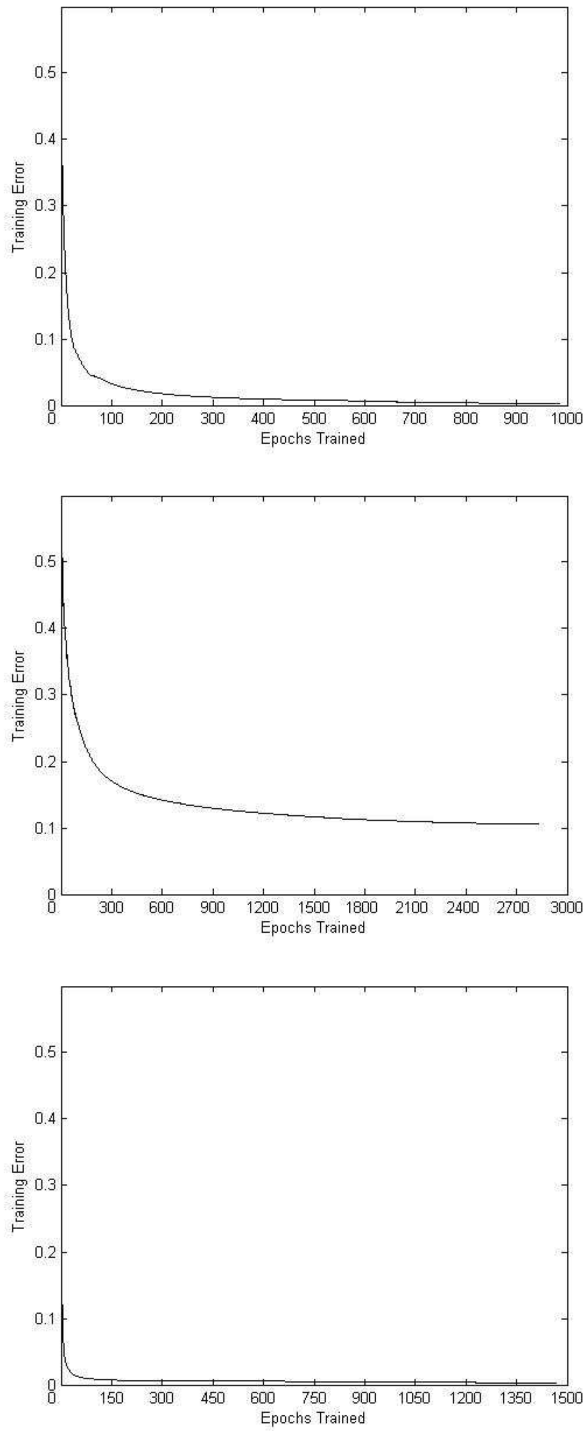

Fig. 3 – Learning curves for the tree learning stages: BA Basic Action (top), EL Entry Level (middle) and HL Higher-Level (bottom).

The first stage reached an average final error of 0.004 at epoch 1000. At the end of the stage, the imitator was able to execute all actions flawlessly, when presented with an object. The overall average error on the final epoch of the Entry-Level stage was 0.03. In particular, for each of the three cycles, the error was 0.05 (linguistic production), 0.04 (linguistic comprehension), and 0.002 (linguistic imitation).

Finally, in the grounding transfer test the agent was requested to perform a new composite action by giving in input only the new action name (HL1) or the new name together with the basic action names. The table below summarizes the testing results, for both versions of the test:

TABLE I
FINAL ERRORS FOR GROUDNING TRANSFER STAGE

	Hug	**Grab**	**Place**	**average**
HL1	0.024	0.019	0.010	0.018
HL2	0.049	0.053	0.035	0.040

Such low errors confirm our hypothesis that previously grounded symbols are transferred to the new behaviors.

IV. Discussion and Conclusions

The model and experiment above concerned the study of the symbol grounding and the symbol grounding transfer in cognitive robotic agents. The positive results of the grounding transfer test demonstrate that it is possible to design autonomous linguistic agents capable of acquiring new grounded concepts. This type of research can provide a solid basis upon which future cognitive robotic projects could be expanded.

The simulation regards a limited lexical and action set, which needs to be scaled up further to demonstrate the potential of such an approach to build useful autonomous robots, such as in service robotics and human-robot interaction and communication systems. The issue of scaling up and combinatorial complexity (CC) in cognitive systems has been recently addressed by Perlovsky [23]. In linguistic systems, CC refers to the hierarchical combinations of bottom-up perceptual and linguistic signals and top-down internal concept-models of objects, scenes and other complex meanings. Perlovsky proposed the Modeling Field Theory (MFT) as a new method for overcoming the exponential growth of CC in computational intelligent techniques currently used in cognitive systems design. MFT uses neural network techniques and fuzzy dynamic logic to avoid CC and computes similarity measures between internal concept-models and the perceptual and linguistic signals. More recently, Perlovsky [3,24] has suggested the use of MFT specifically to model linguistic abilities. By using concept-models with multiple sensorimotor modalities, a MFT system can integrate language-specific signals with other internal cognitive representations.

The proposal to apply MFT in the language domain is highly consistent with the grounded approach to language modeling discussed above. That is, both accounts are based on the strict integration of language and cognition. This permits the design of cognitive systems that are truly able to "understand" the meaning of words being used by autonomously linking the linguistic signals to the internal concept-models of the word constructed during the sensorimotor interaction with the environment. The combination of MFT systems with grounded agent simulations will permit the overcoming of the CC problems currently faced in grounded agent models and scale up the lexicons in terms of high number of lexical entries and syntactic categories. Ongoing simulations are currently looking at the use of MFT for the scaling up of the lexicon and action repertoire of this model [25].

The potential impact of this research for the development of intelligent systems is great, both for cognitive science and for technology. In cognitive science, the area of embodied cognition regards the study of the functioning and organization of cognition in natural and artificial systems. For example, the Higher-Order leaning procedure is inspired by Barsalou's "reenactment" and "mental simulation" mechanism in the perceptual symbol system hypothesis. Barsalou [9] demonstrates that during perceptual experience, association areas in the brain capture bottom-up patterns of activation in sensory-motor areas. Later, in a top-down manner, association areas partially reactivate sensory-motor areas to implement perceptual symbols simulators. A simulation platform like the one used here can be used to test further embodied cognition theories of language, such as Glenberg's [26] action-compatibility effects. In addition, such an approach can be used to study the development and emergence of language in epigenetic robots [15,27,28,29].

For the technological implications of such a project, the model proposed here can be useful in fields such as that of defense systems, service robotics and human-robot interaction. In the area of defense systems, cognitive systems are essential for integrated multi-platform systems capable of sensing and communicating. Such robots can be beneficial in collaborative and distributed tasks such as multi-agent exploration and navigation in unknown terrains. In service and household robotics, future systems will be able to learn language and world understanding from humans, and also to interact with them for entertainment purposes (e.g. [30]). In human-robot communication systems, robots will develop their lexicon through close interaction with their environment and whilst communicating with humans. Such a social learning context can permit a more efficient acquisition of communication capabilities in autonomous robots, as demonstrated in [31].

REFERENCES

[1] A. Cangelosi, G. Bugmann and R. Borisyuk (Eds.), *Modeling Language, Cognition and Action: Proceedings of the 9th Neural Computation and Psychology Workshop*. Singapore: World Scientific, 2005

[2] A. Cangelosi, D. Parisi, "The processing of verbs and nouns in neural networks: Insights from synthetic brain imaging," *Brain and Language*, vol. 89(2), pp. 401-408, 2004

[3] L. Perlovsky, "Integrating language and cognition," *IEEE Connections*, vol. 2(2), pp. 8-13, 2004

[4] T. Riga, A. Cangelosi and A. Greco, "Symbol grounding transfer with hybrid self-organizing/supervised neural networks," in *IJCNN04 International Joint Conference on Neural Networks*. Budapest, July 2004.

[5] S. Harnad, "The Symbol Grounding Problem," *Physica D*, vol. 42, pp. 335-346, 1990.

[6] C.S. Peirce, "Collected papers. Vol. II: Element of logic," in C. Hartshorne & P. Weiss (Eds.), Cambridge, MA: Belknap.

[7] T.W. Deacon, *The Symbolic Species: The Coevolution of Language and Human Brain*, London: Penguin.

[8] A. Cangelosi, "Evolution of communication and language using signals, symbols and words," *IEEE Transactions on Evolutionary Computation*. Vol. 5(2), pp. 93-101, 2001

[9] L.W. Barsalou, "Perceptual symbol systems," *Behavioral and Brain Sciences*, vol. 22, pp. 577-660, 1999.

[10] C. Burgess, and K. Lund, "Modeling parsing constraints with high-dimensional context space," *Language and Cognitive Processes*, vol. 12, pp. 177-210, 1997.

[11] A. Cangelosi, and S. Harnad, "The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language in perceptual categories," *Evolution of Communication*, vol. 4(1), pp. 117-142, 2000.

[12] S. Harnad, S.J. Hanson, and J. Lubin, "Learned categorical perception in neural nets: Implications for symbol grounding," in V. Honavar & L. Uhr (Eds,) *Symbol Processors and Connectionist Network Models in Artificial Intelligence and Cognitive Modelling: Steps Toward Principled Integration*. Academic Press. 2004, pp. 191-206.

[13] G.P. Prince and Y. Demiris, "Editorial: Introduction to the Special Issue on Epigenetic Robotics," *Adaptive Behavior*, vol. 11(2), pp. 75-77, 2003

[14] A. Cangelosi, A. Greco, and S. Harnad, "From robotic toil to symbolic theft: Grounding transfer from entry-level to higher-level categories," *Connection Science*, vol. 12(2), 2000, pp. 143-162

[15] L. Smith and M. Gasser, "The development of embodied cognition: Six lessons from babies," *Artificial Life*, vol. 11 (1-2), pp. 13-30. (Special issue on Embodied and Situated Cognition), 2005.

[16] F. Pulvermuller, *The Neuroscience of Language. On Brain Circuits of Words and Serial Order*. Cambridge: Cambridge University Press, 2003.

[16] J. McClelland, T. Plunkett, and J. Weng, "Special Issue on Autonomous Mental Development," *IEEE Transactions on Evolutionary Computation*, in press

[17] D. Marocco, A. Cangelosi, S. Nolfi, "The emergence of communication is evolutionary robots," *Philosophical Transactions of the Royal Society London – A*, vol. 361, pp. 2397-2421, 2003

[18] S. Nolfi, and D. Floreano, *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. Cambridge, MA: MIT Press/Bradford Books, 2000

[19] A. Cangelosi, "The emergence of language: Neural and adaptive agent models," *Connection Science*, vol. 17(3-4), pp. 185-190, 2005

[20] A. Cangelosi, D. Parisi D. (Eds.) *Simulating the Evolution of Language*. London: Springer-Verlag, 2002

[21] A. Cangelosi, and T. Riga, "An embodied model for sensorimotor grounding and grounding transfer: Experiments with epigenetic robots". *Cognitive science*, in press

[21] L. Steels, "Evolving Grounded Communication for Robots," *Trends in Cognitive Sciences*, vol. 7(7), pp. 308-312, 2003

[22] M. Desmurget, and C. Prablanc, "Postural control of three-dimensional prehension movements," *Journal of Neurophysiology*, vol. 77, pp. 452-464, 1997.

[22] P. Vogt, "Editorial: Language Acquisition and Evolution," *Adaptive Behavior*, vol. 13, pp. 265-268, 2005

[23] L. Perlovsky. *Neural Networks and Intellect: Using Model-Based Concepts*. Oxford University Press, New York, 2001.

[24] J.F. Fontanari, and L. Perlovsky, "Meaning creation adn modeling field theory," in C. Thompson & H. Hexmoor, *Proceedings of KIMAS2005: International Conference on Integration of Knowledge Intensive Multi-Agant Systems*, IEEE Press, 2005, pp. 405-410.

[25] V. Tikhanoff, A. Cangelosi, L. Perlovsky, and J.F. Fontanari, "The integration of language and cognition in robots: Scaling up the lexicon with modeling field theory". *manuscript* in preparation

[26] A. Glenberg, and K. Kaschak, "Grounding language in action," *Psychonomic Bulletin & Review*, vol. 9(3), pp. 558-565, 2002.

[27] P.F. Dominey, “From sensorimotor sequence to grammatical construction: Evidence from simulation and neurophysiology,” *Adaptive Behavior*, vol. 13, pp. 347-361, 2005

[28] B. Macwhinney, “The emergence of linguistic form in time,” *Connection Science*, vol. 17(3-4), pp. 191–211, 2005

[29] L. Steels, “The emergence and evolution of linguistic structure: from lexical to grammatical communication systems,” *Connection Science,* vol. 17(3-4), pp. 213-230, 2005

[30] V. Tikhanoff, and E.R. Miranda, “Musical Composition by an Autonomous Robot: An Approach to AIBO Interaction", in *Proceedings of TAROS 2005 (Towards Autonomous Robotic System)*, Imperial College, London, UK. 2005, pp. 181-188.

[31] L. Steels, and K. Kaplan, “AIBO's first words: The social learning of language and meaning,” *Evolution of Communication*, vol. 4(1), pp. 3-32, 2000.