

Cross-situational learning of object-word mapping using Neural Modeling Fields

José F. Fontanari^{*,a}, Vadim Tikhanoff^b, Angelo Cangelosi^b, Roman Ilin^c,
Leonid I. Perlovsky^{c,d}

^a*Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369,
13560-970 São Carlos SP, Brazil*

^b*Adaptive Behaviour & Cognition Research Group, University of Plymouth, Plymouth
PL4 8AA, UK*

^c*Air Force Research Laboratory, 80 Scott Drive, Hanscom Air Force Base, MA 01731*

^d*Harvard University, 33 Oxford St, Rm 336, Cambridge MA 02138*

Abstract

The issue of how children learn the meaning of words is fundamental to developmental psychology. The recent attempts to develop or evolve efficient communication protocols among interacting robots or virtual agents have brought that issue to a central place in more applied research fields, such as computational linguistics and neural networks, as well. An attractive approach to learning an object-word mapping is the so-called cross-situational learning. This learning scenario is based on the intuitive notion that a learner can determine the meaning of a word by finding something in common across all observed uses of that word. Here we show how the deterministic Neural Modeling Fields (NMF) categorization mechanism can be used by the learner as an efficient algorithm to infer the correct object-word mapping. To achieve that we first reduce the original on-line learning problem to a batch learning problem where the inputs to the NMF mechanism are all possible object-word associations that could be inferred from the cross-situational learning scenario. Since many of those associations are incorrect, they are considered as clutter or noise and discarded automatically by a clutter detector model

*Corresponding author. Fax:+55-16-33739877

Email addresses: fontanari@ifsc.usp.br (José F. Fontanari),
vadim.tikhanoff@plymouth.ac.uk (Vadim Tikhanoff), a.cangelosi@plymouth.ac.uk
(Angelo Cangelosi), roman.ilin.ctr@hanscom.af.mil (Roman Ilin),
Leonid.Perlovsky@hanscom.af.mil (Leonid I. Perlovsky)

included in our NMF implementation. With these two key ingredients – batch learning and clutter detection – the NMF mechanism was capable to infer perfectly the correct object-word mapping.

Key words: Cross-situational learning; Language acquisition; Clustering algorithms; Neural Modeling Fields

1. Introduction

Computational models play an important role in the investigation of language, both from an evolutionary perspective and from a developmental timescale. For example, simulation and robotic models of the evolution of language (Parisi & Cangelosi , 2002; Kirby , 2002) provide a complementary methodology that can help researchers to develop detailed and precise hypotheses on language origins and evolution, and to test these hypotheses in the virtual experimental laboratory of the simulation. Other models focus on the study of the developmental stages in language acquisition. This is the case of epigenetic robotics, where the design of the robot’s behavioral and cognitive capabilities is directly inspired by developmental psychology theories (Lungarella et al. , 2003; Cangelosi & Riga , 2006). These can, for example, implement well-known phenomena in language development, such as vocabulary spurt and mutual exclusivity (Tomasello , 2003).

In this contribution we carry further the rather ambitious research program of integrating language and cognition within the Neural Modeling Fields Framework (NMF) (Perlovsky , 2004, 2006a; Tikhonoff et al. , 2006; Cangelosi et al. , 2007). This is a task of enormous breadth that encompasses many unsolved (and, perhaps, unsolvable) problems such as object perception (Marr , 1982; Kellman & Spelke , 1983), symbol grounding which addresses the question of how physical signs can be given meaning (Harnard , 1990; Cangelosi et al. , 2002), and the emergence of a common lexicon in a population of interacting agents (Cangelosi , 2001; Fontanari & Perlovsky , 2007, 2008a). This unifying perspective accords with the view that language is not an isolated capability of the individual and cannot be fully comprehended if one ignores its intrinsic relationships with the cognitive and social abilities (Cangelosi et al. , 2002).

It must be noted, however, that this integrated view of language is not unanimously accepted by the linguist community. In fact, Bickerton (1990) argues that language is primarily a representational system, established well

before our remote ancestors have uttered the first recognizable word. In the same vein, Chomsky (2000) claims that "...language is not properly regarded as a system of communication. It is a system for expressing thought, something quite different." Both views purport that language is not for communication, rather its main use being inner-speech for adults and monologue for children. Of course, this stance collides with the conventional evolutionary vision of language, according to which language evolved under a selective pressure towards better communication (Pinker & Bloom, 1990).

Although, as illustrated above, almost every statement about language can be considered controversial, one fact is noncontroversial, namely, that the lexicon must be learned from the active or passive interaction between children and language-proficient adults. Of course, the issue of whether this ability to learn the lexicon is due to some domain-general learning mechanism or is an innate ability, unique to humans, is still on the table (Bates & Elman, 1996). There is, nonetheless, a general agreement that lexicon learning is accomplished by children chiefly through unsupervised learning (Bloom, 2000).

Here we explore one possible unsupervised learning scenario for lexicon acquisition, namely, the cross-situational learning scenario, which is based on the very intuitive idea that one way that a learner can determine the meaning of a word is to find something in common across all observed uses of that word (Siskind, 1996). Although, the general notion of cross-situational learning has been proposed by many authors (see, e.g., Pinker (1989)) the instantiation of this rather abstract notion into a manageable mathematical or computational model is still subject of intense research (Smith, 2003; Lenaerts et al., 2005; de Beule et al., 2006).

Essentially, language can be reduced to a mapping between sounds and meanings – the tenet of Chomsky’s Minimalist Program (Chomsky, 1995). Here we address a part of this problem, in which language is viewed as a mapping between sounds (or words) and objects in the world. This is obviously a necessary first step to tackle more realistic problems, where the words may refer to particular sets of objects rather than to single objects only. In particular, we equip the learner with a NMF categorization mechanism sensitive to the frequency of co-occurrence of objects and words and show how the inclusion of a clutter detection module can identify and automatically discard the inputs representing incorrect object-word associations. This is a major add-on to our previous attempt to apply the NMF categorization mechanism to the cross-situational lexicon acquisition problem (Fontanari et

al. , 2009).

The rest of the paper is organized as follows. In Section 2 we present the cross-situational learning scenario and describe the structure of the input set for the batch learning mode. In Section 3 we describe at some length the NMF categorization mechanism within the context of the specific problem addressed in this paper. In Section 4 we present the results of the simulations of the NMF dynamics. Finally, in Section 5 we summarize our main conclusions and discuss an alternative NMF formulation for the cross-situational learning problem. A previous attempt to address the cross-situational learning of object-word mapping using the NMF algorithm was published in Fontanari et al. (2009).

2. Cross-situational learning scenario

Our minimal model of cross-situational lexicon learning involves two agents: the teacher, who has complete domain of the language, described by a one-to-one object-word mapping, and the pupil, equipped with a ‘NMF-mind’, and who has no knowledge of object-word relations. The pupil must infer that mapping from the examples provided by the teacher.

We assume that there are N objects in the agents’ world. At each learning event, two objects are chosen randomly without replacement to form the context and the teacher then names one of them. The pupil has access to the context as well as to the word emitted by the teacher. The pupil’s task is to guess to which object in the context the word refers to. Figure 1 illustrates the cross-situational lexicon acquisition scenario.

The deterministic nature of the NMF algorithm requires that all input data (e.g., all possible examples of object-word pairs) be presented at once to the pupil. This sort of batch-mode learning setting differs considerably from the on-line learning of the typical guessing game scenario. Of course, real life learning is sequential as children are exposed to situations one at a time. However, learning is cumulative: a child looks around for several years and then very fast he/she learns to understand many situations. To simplify, we assume here that the data about all the situations are available at the same time. This assumption can be modified in the future.

For the sake of concreteness, let us label objects and words by the integers $1, 2, \dots, N$, and assume without lack of generality that the correct one-to-one object-word mapping is such that object 1 corresponds to word 1, object 2 to word 2, and so on. A learning event consists of the presentation

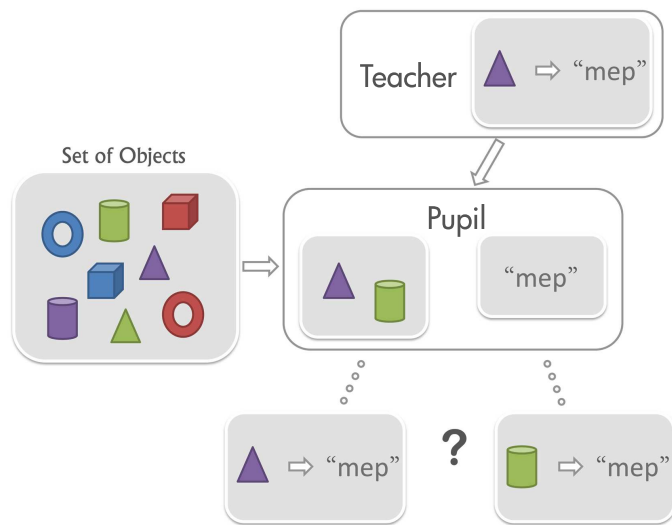


Figure 1: Scheme of the cross-situational lexicon acquisition scenario. The pupil has access to a context comprised of two objects selected at random from the full set of N objects and to a word ('mep' in the figure) uttered by the teacher. The pupil's task is to infer the word-object mapping from the co-occurrence of objects in the context and the accompanying words.

of one of the following $N(N - 1)$ situations: $\{(1, 1), (2, 1)\}$, $\{(1, 2), (2, 2)\}$, $\{(1, 1), (3, 1)\}$, \dots , $\{(N, N - 1), (N - 1, N - 1)\}$. Here the first brace represents the situation where the context comprises objects 1 and 2 and the teacher utters the word that corresponds to object 1; the second brace represent the same context but now the accompanying word corresponds to object 2, etc. We have separated the input data by braces which represent different learning events for the sake of clarity only: because of the batch-mode of learning, the pupil has no access to that kind of information. Note that because the two objects that form the context are chosen without replacement the N situations $\{(1, 1), (1, 1)\}$, dots, $\{(N, N), (N, N)\}$ can never occur. Hence the total number of learning situations (i.e., of braces in the above notation) is $N^2 - N = N(N - 1)$. Since each situation consists of two object-word associations, the total number of inputs to the pupil is $M \equiv 2N(N - 1)$. Clearly, the pair $(1, 1)$ appears $N - 1$ times, whereas the pair $(1, n); n = 2, \dots, N$ appears just once, and similarly for the other objects. So the number of correct associations (n, n) is $N(N - 1)$. The same result holds for the number of incorrect associations $(n, m); n \neq m$ as well.

The pupil attempts to model the input data through the object-word mapping (S_{1k}, S_{2k}) , with $k = 0, \dots, N$, where the components $S_{ek}, e = 1, 2$ – the so-called modeling fields – are real variables given by the NMF equations described in the next section. The question is whether the pupil can recover the correct object-word mapping $(1, 1), (2, 2), \dots, (N, N)$ having access only to the information available in the cross-situational learning scenario summarized in Fig. 1. We note that there is a considerable loss of information involved in the procedure of replacing the sequential presentations of contexts plus words by the entire set of the M object-word examples, since one can easily imagine different situations which result in the same batch-mode learning set.

3. Neural Modeling Fields dynamics

The Neural Modeling Fields (NMF) algorithm proposed by Perlovsky (2001) is essentially an iterative, self-consistent, deterministic process designed to maximize the similarity between models and incoming signals. In this aspect, it shares some elements with the Hopfield-Tank neural network (Hopfield & Tank, 1985) and the mean-field annealing (Bilbro et al., 1989). In fact, these two deterministic heuristics have been extensively used to search for optimal or quasi-optimal solutions of a variety of optimization

problems, whereas NMF searches for the maximum of a global similarity function. The main feature that sets NMF apart from these heuristics, as well as from many other neural networks, is use of parametric models. The so-called fuzzy association variables, which can be thought of as weights of a neural network (Perlovsky & McManus, 1991) are expressed as functions of these models. These variables give a measure of the probability of association between input data and concept-models, although, as already pointed out, NMF is a deterministic algorithm.

As discussed in Section 2, each example (input signal) is described by the pair of integer variables (O_{1i}, O_{2i}) with $i = 1, \dots, M$. In most applications of the NMF algorithm, the aim is to maximize the similarity between input data and parametric models used to represent (and usually compress) that data. The difficulty with the cross-situational learning scenario is that some data must be ignored, as they represent wrong object-word associations. Let us assume that there are N models described by the pairs (S_{1k}, S_{2k}) with $k = 1, \dots, N$ that should ‘model’ the original set of M examples; hence the denomination ‘modeling fields’ to the mathematical quantities S_{ek} . In addition to these regular object-word pair models, we introduce a trash-can model $k = 0$ which is described by a constant value. Finally, to each model k we associate a *a priori* probability r_k such that $\sum_{k=0}^N r_k = 1$.

We begin the derivation of NMF algorithm, i.e., of the equations that govern the dynamics of the modeling fields, by introducing a measure for the similarity between input $i = 1, \dots, M$ and model $k = 1, \dots, N$, namely,

$$l(i | k) = \prod_{e=1}^2 [2\pi\sigma_{ke}^2]^{-1/2} \exp \left[-\frac{(O_{ie} - S_{ke})^2}{2\sigma_{ke}^2} \right] \quad (1)$$

where, at this stage, the fuzziness σ_{ke}^2 are parameters given *a priori*. For model $k = 0$, we define

$$l(i | 0) = 1/M \quad (2)$$

for all i . Clearly, model $k = 0$ plays the role of a trash can or clutter detector: it will collect any input i that is not close enough to one of the regular models. The goal is to find an assignment between models k and inputs i such that the global similarity

$$L_0 = \sum_{i=1}^M \ln \sum_{k=0}^N r_k l(i | k) \quad (3)$$

is maximized with respect to the modeling fields S_{ek} and prior probabilities r_k . To take the normalization of r_k into account we maximize the quantity

$$L_\lambda = \sum_{i=1}^M \ln \sum_{k=0}^N r_k l(i | k) + \lambda \left(\sum_{k=0}^N r_k - 1 \right) \quad (4)$$

where λ is a Lagrange multiplier. Since

$$\frac{dL_\lambda}{dt} = \sum_{k=1}^N \sum_{e=1}^2 \frac{\partial L_\lambda}{\partial S_{ke}} \frac{dS_{ke}}{dt} + \sum_{k=0}^N \frac{\partial L_\lambda}{\partial r_k} \frac{dr_k}{dt} \quad (5)$$

we can guarantee that $dL_\lambda/dt \geq 0$ by defining the dynamics $dS_{ke}/dt = \partial L/\partial S_{ke}$ and choosing λ such that $\partial L/\partial r_k = 0$.

The calculation of $\partial L_\lambda/\partial S_{ke}$ is straightforward and yields

$$\frac{\partial L_\lambda}{\partial S_{ke}} = \sum_{i=1}^M \frac{r_k}{\sum_{k'=0}^N r_{k'} l(i | k')} \frac{\partial l(i | k)}{\partial S_{ke}}. \quad (6)$$

Using the identity $\partial y/\partial x = y \partial \ln y/\partial x$ and defining the fuzzy associations as

$$f(k | i) = \frac{r_k l(i | k)}{\sum_{k'=0}^N r_{k'} l(i | k')} \quad (7)$$

we can immediately write the equations for the modeling fields (Perlovsky , 2001)

$$\frac{dS_{ke}}{dt} = \sum_{i=1}^M f(k | i) \frac{\partial \ln l(i | k)}{\partial S_{ke}} \quad (8)$$

for $k = 1, \dots, N$ and $e = 1, 2$.

The condition $\partial L_\lambda/\partial r_k = 0$ is written as

$$\sum_{i=1}^M \frac{l(i | k)}{\sum_{k'=0}^N r_{k'} l(i | k')} + \lambda = 0. \quad (9)$$

Multiplying by r_k and summing over k yields $\lambda = -M$ so that

$$r_k = \frac{1}{M} \sum_{i=1}^M f(k | i). \quad (10)$$

We note that given $l(i | k)$ [see Eqs. (1) and (2)] we can obtain r_k by solving Eqs. (7) and (10) simultaneously. Hence the r.h.s. of our fundamental dynamic equation (8) is a function of the modeling fields S_{ek} only, as expected. In practice, the exact determination of r_k is not important provided they are approximately the same for all models. In our analysis we will fix r_k to the values corresponding to the optimal solution of the learning problem (see Section 4).

The fuzzy associations variables $f(k | i)$, Eq. (7), play a fundamental role in the interpretation of the NMF dynamics by giving a measure of the correspondence between input i and model k relative to all other models k' . We note that the factor $f(k | i)$ in Eq. (8) couples not only S_{1k} and S_{2k} which is critical for producing a sensible object-word mapping but also the components of different modeling fields.

By construction, the dynamics (8) always converges to a (possibly local) maximum of the similarity L_λ for fixed fuzziness σ_{ke}^2 . A salient feature of the NMF is a match between parameter uncertainty and fuzziness of similarity. By properly decreasing the value of the fuzziness σ_{ke}^2 , a unique assignment between inputs and models is attained. In fact, for fixed σ_{ke}^2 we obtain the fuzzy logic limit, whereas for $\sigma_{ke}^2 = 0$ we obtain the usual crispy, Aristotelian logic limit. The basic idea of NMF is to reduce the fuzziness during the time evolution of the modeling fields and so, because of this interpolation, the algorithm is also referred as Dynamic Logic. Of course, this procedure is similar to the cooling schedule of simulated annealing (Kirkpatrick et al. , 1983), although, as pointed out before, a more appropriate comparison is with mean-field annealing (Bilbro et al. , 1989) since both algorithms are deterministic.

In what follows we decrease the fuzziness on the fly, i.e., simultaneously with the change of the modeling fields, according to the following prescription (Fontanari & Perlovsky , 2005)

$$\sigma_{ke}^2(t) = a_{ke}^2 + b_{ke} \exp(-\alpha_e t) \quad (11)$$

where $\alpha_e = 10$, $a_{ke} = 0.2$, and $b_{ke} = 1.0$ are time-independent control parameters. As a guideline for setting the values of these parameters, we note that b_{ke} must be chosen large enough such that, at the beginning, all category examples can be described by all modeling fields, whereas the baseline resolution a_{ke} must be small enough such that, at the end, a given modeling field will describe a single category. However, a_{ke} should not be set to a too

small value to avoid numerical instabilities in the calculation of the partial similarities defined by Eq. (1).

4. Results

It is easy to characterize the optimal performance: the $N - 1$ identical inputs $(1, 1)$ must be associated to model $k_1 > 0$, the $N - 1$ identical inputs $(2, 2)$ to model $k_2 > 0$, etc. , and the remaining $N(N - 1)$ inputs corresponding to wrong object-word associations must be associated to model $k = 0$. In terms of the fuzzy associations, the optimal solution is $f(k_1 | i) = 1$ if input i is the pair $(1, 1)$, and 0 otherwise; $f(k_2 | i) = 1$ if input i is the pair $(2, 2)$, and 0 otherwise; etc., and $f(0 | i) = 1$ if input i is a wrong object-word association, and 0 if i is a correct object-word association. Hence, this optimal solution corresponds to the prior probabilities [see Eq. (10)]: $r_k = 1/2N$ for $k = 1, \dots, N$ and $r_0 = 1/2$. The results of this section were obtained by solving the ordinary differential equations (8) using these values for the prior probabilities.

In Fig. 2 we show the results of a typical run in which the initial condition (i.e., the values of the modeling fields at $t = 0$) are random perturbations of the correct object-word mapping, $(n, n); n = 1, \dots, N$. Explicitly, for model $k = 1, \dots, N$ we set $\vec{S}_k = (k + \epsilon, k + \epsilon)$ where ϵ is a random variable uniformly distributed in the interval $[-1, 1]$. In addition, without lack of generality, we set the inputs from $i = 1$ to $i = N - 1$ to $(1, 1)$; from $i = N$ to $i = 2N - 2$ to $(2, 2)$; ...; from $i = N(N - 1) - (N - 2)$ to $i = N(N - 1)$ to (N, N) . Finally, the wrong object-word associations are distributed among inputs $i = N(N - 1) + 1$ and $i = M = 2N(N - 1)$. The results in this figure show that the NMF algorithm found the optimal solution¹.

We note that panel (a) of Fig. 2 reveals an extraneous asymmetry between the different models: whereas models $k = 1$ and $k = 5$ are already well adjusted to their input subsets, the status of models $k = 2, 3$ and 4 are completely uncertain at the initial stage. This is so because, say, the model $k = 3$ become object of competition between the modeling fields $k = 2, 3$, and 4, whereas there are much less competition for models $k = 1$ and 5.

¹This is a great improvement compared to our previous results (Fontanari et al. , 2009). The successful performance owns to the inclusion of the trash-can model $k = 0$, which absorbed all the wrong object-word associations.

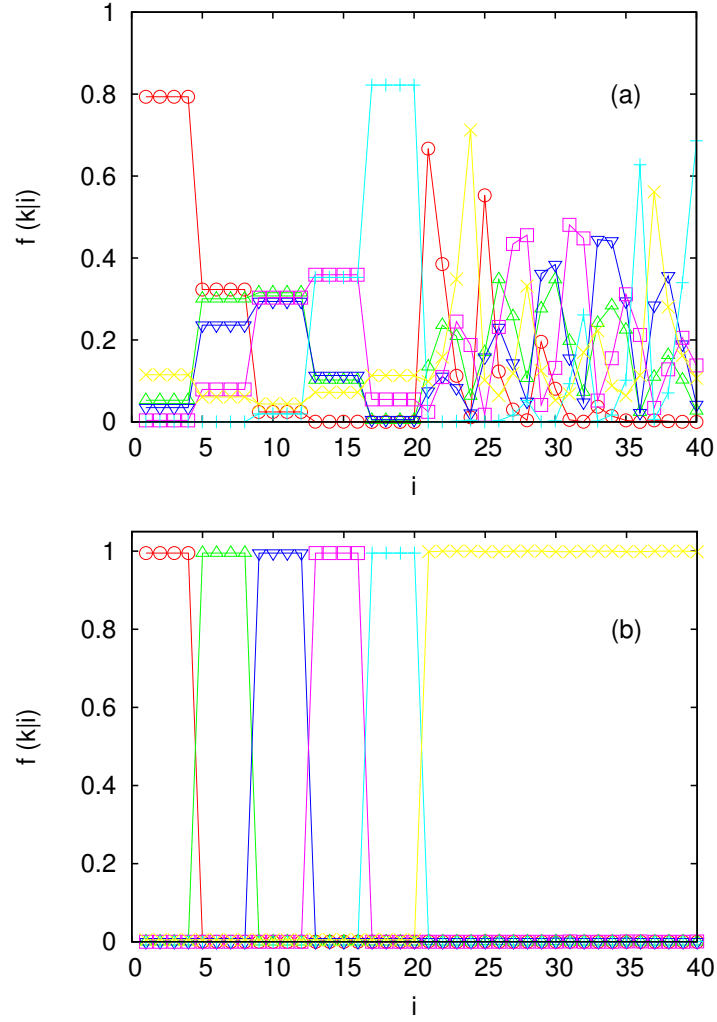


Figure 2: The fuzziness association variables $f(k|i)$ for inputs $i = 1, \dots, 40$ in the case of $N = 5$ objects. The inputs are rearranged as described in the text and the modeling fields are initialized as random perturbations of the correct object-word mapping. The convention is $k = 0 (\times)$, $1 (\circ)$, $2 (\triangle)$, $3 (\nabla)$, $k = 4 (\square)$, and $k = 5 (+)$. Panels (a) and (b) illustrate the initial ($t = 0$) and final ($t \rightarrow \infty$) situations, respectively.

5. Conclusion

Interaction between language and cognition is among unknown neural mechanisms of the mind and a problem posing significant mathematical difficulties. These mechanisms have to explain several aspects of the mind functioning, which seem mysterious. How does language acquire meanings? Is cognition based on language, or vice-versa? Do we think with words, or do we use words to label already made up thoughts and decisions, and if so, then what are thoughts without words? Why children acquire language by the age of 5, but cannot act like adults, what exactly is missing in terms of neural mechanisms? How is it possible that we learn correct associations between words and objects, phrases and situations, among astronomically large number of possible incorrect associations? How phrases are formed from words so that they relate to the meanings of situations in the real world? An outline of possible mathematical and neural mechanisms has been described in Perlovsky (2006b, 2007a,b, 2009a).

The current contribution along with Fontanari & Perlovsky (2005, 2007, 2008a,b) is a significant step in the mathematical implementation of that program. We have demonstrated lexicon acquisition in a cross-situation learning using NMF. The example demonstrated here is intentionally limited in scope and can be solved by mere statistical means. Its significance, however, is in that NMF is a scalable mathematical paradigm, its complexity grows linearly with the number of words and objects, and it has been demonstrated to solve problems, which could not have been solved by other methods due to combinatorial complexity (Perlovsky, 2006b, 2009a).

Future directions include multi-agent simulations (Fontanari & Cangelosi, 2009), in which models are replaced by competing agents; mathematical methods implementing and demonstrating learning of phrases from words, situations from objects (Ilin et al., 2009), and extending these methods to abstract thinking in interacting language and cognition. These mathematical methods will be combined with perceptual symbol system (Barsalou, 1999) and related to neural mechanisms and brain imaging studies. Future research will address the role of language in rational and irrational thinking (Levine & Perlovsky, 2008); emotional mechanisms of language and roles of language emotionality and music in interaction of language and cognition (Perlovsky, 2008, 2009a); co-evolution of languages, cognition, and cultures (Perlovsky, 2007c, 2009b).

Acknowledgments

The research at São Carlos was supported in part by CNPq and FAPESP, Project No. 04/06156-3. AFRL research was supported in part by AFOSR PM Dr. D. Cochran, and by the NRC Research Associateships programs award.

References

- Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577-609.
- Bates, E., & Elman, J. (1996). Learning rediscovered. *Science*, *274*, 1849-1850.
- de Beule, J., de Vylder, B., & Belpaeme, T. (2006). A cross-situational learning algorithm for damping homonymy in the guessing game. In: L.M. Rocha, M. Bedau, D. Floreano, R. Goldstone, A. Vespignani, L. Yaeger (Eds.) *Proceedings of the Xth Conference on Artificial Life* (pp. 466-472). Cambridge, MA: MIT Press.
- Bickerton, D. (1990). *Language & Species*. Chicago: University of Chicago Press.
- Bilbro, G., Mann, R., & Miller, T. (1989). Optimization by mean field annealing. In: D. S. Touretzky (Ed.), *Advances in Neural Information Processing Systems I* (pp. 91-98). San Mateo, CA: Morgan Kaufmann.
- Bloom, P. (2000). *How children learn the meaning of words*. Cambridge, MA: MIT Press.
- Cangelosi, A. (2001). Evolution of Communication and Language Using Signals, Symbols, and Words. *IEEE Transactions on Evolutionary Computation*, *5*, 93-101.
- Cangelosi, A., Greco, A., & Harnard, S. (2002). Symbol grounding and the symbolic theft hypothesis. In: A. Cangelosi, & D. Parisi (Eds.), *Simulating the Evolution of Language* (pp. 191-210). London: Springer-Verlag.
- Cangelosi, A., & Riga, T. (2006). An embodied model for sensorimotor grounding and grounding transfer: Experiments with epigenetic robots, *Cognitive Science*, *30*, 673-689.

- Cangelosi, A., Tikhanoff, V., Fontanari, J. F., & Hourdakis E. (2007). Integrating Language and Cognition: A Cognitive Robotics Approach, *IEEE Computational Intelligence Magazine*, 2, 65-70.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: Cambridge University Press.
- Chomsky, N. (2000). *On nature and language*. New York: Cambridge University Press.
- Fontanari, J.F., & Perlovsky, L. I. (2005). Meaning Creation and Modeling Field Theory. *Proceedings of the IEEE Conference on Integration of Knowledge Intensive Multi-Agent Systems* (pp. 405-410).
- Fontanari, J.F., & Perlovsky, L.I. (2007). Evolving compositionality in evolutionary language games. *IEEE Transactions on Evolutionary Computation*, 11, 758-769.
- Fontanari, J.F., & Perlovsky, L.I. (2008a). A game theoretical approach to the evolution of structured communication codes. *Theory in Biosciences*, 27, 205-214.
- Fontanari, J.F., & Perlovsky, L.I. (2008b). How language can help discrimination in the Neural Modeling Fields framework. *Neural Networks*, 21, 250-256.
- Fontanari, J.F., Tikhanoff, V., Cangelosi, A., & Perlovsky, L.I. (2009). A cross-situational algorithm for learning a lexicon using Neural Modeling Fields. *Proceedings of the IEEE International Joint Conference on Neural Networks IJCNN09*.
- Fontanari, J.F., & Cangelosi, A. (2009). A minimal model of cross-situational vocabulary learning. *Proceedings of the IEEE Conference on Integration of Knowledge Intensive Multi-Agent Systems*.
- Harnard, S. (1990). The symbol grounding problem. *Physica D*, 42, 335-346.
- Hopfield, J.J., & Tank, D.W. (1985). Neural computation of decisions in optimization problems. *Biological Cybernetics*, 52, 141-152.
- Ilin, R., Fontanari, J. F., & Perlovsky, L. I. (2009). Learning situations with multiple objects. Unpublished.

- Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, *15*, 483-524.
- Kirby, S. (2002). Natural language from artificial life. *Artificial Life*, *8*, 185-215.
- Kirkpatrick, S., Gelatt, C.D., & Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*, *220*, 671-680.
- Lenaerts, T., Jansen, B., Tuyls, K., & de Vylder, B. (2005). The evolutionary language game: An orthogonal approach. *J. Theor. Biol.*, *235*, 566-582.
- Levine, D.S., & Perlovsky, L.I. (2008). Neuroscientific Insights on Biblical Myths: Simplifying Heuristics versus Careful Thinking: Scientific Analysis of Millennial Spiritual Issues. *Zygon, Journal of Science and Religion*, *43*, 797-821.
- Lungarella M., Metta G., Pfeifer R., & Sandini G. (2003). Developmental robotics: A survey. *Connection Science*, *15*, 151-190.
- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- Parisi, D., & Cangelosi, A. (2002). *Simulating the Evolution of Language*. London: Springer-Verlag.
- Perlovsky, L.I., & McManus, M.M. (1991). Maximum Likelihood Neural Networks for Sensor Fusion and Adaptive Classification. *Neural Networks* *4*, 89-102.
- Perlovsky, L. I. (2001). *Neural Networks and Intellect: Using Model-Based Concepts*. Oxford: Oxford University Press.
- Perlovsky, L. I. (2004). Integrating language and cognition. *IEEE Connections*, *2*, 8-13.
- Perlovsky, L. I. (2006a). Integrated Emotions, Cognition, and Language. *Proceedings of the IEEE International Joint Conference on Neural Networks IJCNN06* (pp. 1570 -1575).
- Perlovsky, L.I. (2006b). Toward Physics of the Mind: Concepts, Emotions, Consciousness, and Symbols. *Phys. Life Rev.*, *3*, 22-55.

- Perlovsky, L.I. (2007a). Symbols: Integrated Cognition and Language. In: R. Gudwin, & J. Queiroz (Eds.), *Chapter in Semiotics and Intelligent Systems Development* (pp.121-151). Hershey, PA: Idea Group.
- Perlovsky, L.I. (2007b). Neural Networks, Fuzzy Models and Dynamic Logic. In: R. Köhler, & A. Mehler (Eds.), *Chapter in Aspects of Automatic Text Analysis* (Festschrift in Honor of Burghard Rieger) (pp. 363-386). Berlin: Springer.
- Perlovsky, L.I. (2007c). Evolution of Languages, Consciousness, and Cultures. *IEEE Computational Intelligence Magazine*, 2, 25-39.
- Perlovsky, L.I. (2008). Music and Consciousness. *Leonardo, Journal of Arts, Sciences and Technology*, 41, 420-421.
- Perlovsky, L.I. (2009a). Language and Cognition. *Neural Networks*, doi:10.1016/j.neunet.2009.03.007
- Perlovsky, L.I. (2009b). Emotions, Language, and Sapir-Whorf Hypothesis. *Neural Networks*, this issue.
- Pinker, S. (1989). *Learnability and cognition*. Cambridge, MA: MIT Press.
- Pinker, S. (1994). *The Language Instinct*. London: The Penguin Press.
- Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13, 707-784.
- Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61, 39-91.
- Smith, A.D.M. (2003). Semantic generalization and the inference of meaning. *Lecture Notes in Artificial Intelligence*, 2801, 499-506.
- Tikhanoff, V., Fontanari, J.F., Cangelosi, A., & Perlovsky, L.I. (2006). Language and Cognition Integration Through Modeling Field Theory: Category Formation for Symbol Grounding. *Lecture Notes in Computer Science*, 4131, 376-385.
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.