COMPLEXITY INTERNATIONAL

Volume 09

http://www.csu.edu.au/ci/

Mutual influence of learning and evolution

H. J. Dopazo [1,(*)] and R. P. J. Perazzo [2]

[1] *Lab. Biología del Comportamiento. Instituto de Biología y Medicina Experimental (IByME-CONICET). Buenos Aires, Argentina.*

[(*)] *Present Address: Unidad de Bioinformática. Centro Nacional de Investigaciones Oncológicas (CNIO). c/ Melchor Fernández Almagro 3, 28029, Madrid, España*

Email: hdopazo@cnio.es

[2] *Centro de Estudios Avanzados (CEA), Depto. de Física. (Facultad de Ciencias Exactas y Naturales (FCEyN). Universidad de Buenos Aires (UBA). Ciudad Universitaria, 1425, Buenos Aires, Argentina.*

Email: rperazzo@df.fcen.uba.ar

Abstract

We review several models that deal with the interaction of the genetic and behavioral systems throughout evolution. We present the conjecture of Baldwin and the subsequent developments of his ideas and those developed by Schmalhausen. We discuss the experiments of genetic assimilation of Waddington and the modern viewpoint based in the increasingly important role of the modification that the individuals do of the environment in which they dwell. Finally we consider the theoretical model of Hinton & Nowlan and an extension that involves a population of adaptive perceptrons in which some of their synapses can be updated through a learning process. This new approach leads to a halting of the Baldwin Effect due to efficient learning. This causes that the transcription of environmental data into genetic information remains hindered by learning, instead of stimulated as it is usually understood. From the perspective of the general description of complex systems, the interaction of learning and evolution can be considered as the mutual influence between two different hierarchical levels of selection taking place in two widely different timescales.

1. Introduction

Survival and reproduction of living beings depends upon the access to scarce resources. These are in addition distributed irregularly in space and in time. As a consequence learning or what is the same, the capacity to change the behavior on the basis of past experiences can not occur arbitrarily. All the important features of living beings must contribute positively to their reproductive success and are transferred with almost no change to the descendants through inheritance. The ancestors of all beings that live today have been successful in managing that diffuse information, at least until they have reached reproductive maturity.

Received: 08 Oct 2002
Accepted: 08 Oct 2002

In this fashion, natural selection and historical background are the true architects of the natural world. This has been the central idea of the transcendental work of Charles Darwin in 1859, ‘‘The origin of species by means of natural selection"that has framed all modern biological thinking. From a general point of view, learning capability has evolved to soften the effects produced by changeable and unpredictable surroundings. In the extreme opposite case, the most economic evolutionary strategy for an organism developing in a fixed environment, is to get rid of the learning capability and encode its behavior in some portion of its genetic information. Such situation is however far from any real circumstance.

All organisms, from the most simple to the most complex, have means to assess the conditions of the environment where they live and provide an adaptive response to it. Bacteria such as *E. coli* have developed biochemical sensing systems to regulate its displacements. Through their membrane proteins,it can detect nutrients and toxins sending information to the interior of the cell and modify the direction of its displacement accordingly. In a similar fashion the brain receives information about the possible presence of dangers or of resources. That information is stored and assessed triggering complex adaptive motor responses that shape the conduct of the individual.

It is evident that biological evolution has developed the learning systems to provide effective responses to the challenges of the environment. What is far from being evident is if the inverse process has any biological relevance i.e. that learning has any effects over the evolutionary process, or equivalently, if everyday actions of the living organisms can guide the evolution of the whole population and, consequently that of the species. If such mutual influence does indeed exist, one could also be able to detectfeatures of living organisms that can only be explained through a process of this kind. In the following sections we discuss these questions. We first review the background and we later present the results of theoretical models that have been developed to discuss this effect.

2. Background

The possibility that acquired characters could be transmitted to the descendants has been a recurrent idea in Biology. The French scientist J. B. Lamarck (1744-1829) was the first to suggest a biological evolutionary theory in which the progressive change was precisely based upon this type of transmission. Darwin did not reject these ideas although he thought that selection was the primary driving force to produce cumulative changes. August Weissmann, a practicing evolutionary cytologist and embryologist, was the one who completely discarded that idea because he could not conceive a mechanism to encode into the gametes features such as a fully developed muscle or the expertise to find water wells. The present development of modern genetics as well as molecular biology has provided further support to reject the idea that an organic adaptation can change the information encoded in the gametes. The consequence of this is that many natural scientists still consider natural selection as an external force acting upon organisms that are passive in their relationship with the environment.

One important achievement of the early evolutionary theory is that a clear, conceptual distinction was made between the ‘‘organism" and the ‘‘environment", but, in doing so, even Darwin denied the active role of living beings in changing each other and their surrounding environment. For Darwin and for all who have produced what is known as the ‘‘Evolutionary Synthesis", the environment is only the agent that takes care of the process of selection, or, to put it into different words, is the arena where phenotypes show different survival or reproductive successes. Darwin said:

> *...natural selection is daily and hourly scrutinizing, throughout the world, the slightest variations; rejecting those that are bad, preserving and adding up all that are good; silently and insensibly working, whenever and wherever opportunity offers, at the improvement of each organic beingin relation to its organic and inorganic conditions of life.*

The result of this line of thought was to conceive the environment as a static template where natural selection moulds all living organisms by changing its functional features and thus insuring their reproductive success. Thus the evolutionary theory that derives from the Evolutionary Synthesis extended the model of substitution of extant allelic variations as the only and exclusive origin of adaptation.

At the end of the past century H. F. Osborn (1896), C. Lloyd Morgan (1896) and J. M. Baldwin (1896) suggested a different mechanism for the origin of adaptations (see Gotltieb (1992) and Bellew & Mitchell, (1996) for an historical perspective of the issue). Organisms where no longer considered as passive entities facing a fixed environment. Instead, they where assumed to change their surroundings through their behavior or other features that could *not* be transmitted through inheritance. In turn, that same environment that has been changed by the individuals, plays the role of being the main responsible for their selection.

In this way the survival - and not the death - of the living beings in an altered environment could favor the assimilation, through natural selection, of genetic mutations that may be helpful for the adaptation to that new environment. The process of physiological adaptation is thus considered to be the basis for the assimilationof favorable genetic variations. Baldwin called this process one of ‘‘organic selection" and suggested that could help evolution providing (or facilitating) a path through which natural selection - based uponcumulative genetic changes - could drag the population of individuals towards an optimum in adaptation. He suggested that this is the way in which what had to be learned in previous generations could appear to be genetically encoded in later ones. In spite of the seemingly Lamarckian content of his statements, Baldwin provided the arguments justifying that the process could take place according within a pure Darwinian framework.

G. G. Simpson (1953), a paleontologist that is one of the founder of the Evolutionary Synthesis christened this process as the "Baldwin effect".

Previous to that, the geneticist I. Schmalhausen made a point in regarding this selection process only as a result of the interaction of the phenotype with the environment during the development of the individual (see discusion in Schlichting & Pigliucci, 1998). His concept was that behavioral traits alter the physical and social environment, modifying in turn the selection coefficients that guide the changes between generations. He emphasized that characters are not rigidly determined by the genotype claiming instead that this fixes the range for developing a given character. Environmentally induced "acquired characters" can occur within that range, but is the natural selection that acts upon the genetic system the one which modifies the range of phenotypic responses.

Schmalhausen thought that evolution primarily changed the ‘‘reaction norm" of the genotype. Such reaction norm is the environmental modulation of the mapping between genotype and phenotypes, i.e. how can the same genotype express itself through different phenotypes in different environments. Adaptive evolution takes place when environmental changes exhibit a previously unexposed portion of the reaction norm. This step is followed by selection of mutations that can improve adaptations in the new environment. This was considered by him a special case of ‘‘stabilizing selection". This evolutionary mechanism is perhaps more noticeable when dealing with characters acquired through learning because these can very rapidly be transmitted across generations by cultural means.

More recently C. H. Waddington (1942) proved experimentally that a phenotypic character that is only expressed by means of the presence of an environmental stimulus, could later be fixed into the genome, through a selective process, consequently at the end of his experiment the trait turned out to be independent of the existence of the environmental stimulus. Waddington called this process one of ‘‘genetic assimilation" and said that ‘‘...acquired characters in the conventional sense could be assimilated by the genotype, and eventually appear comparatively independent of any specific environmental influence". This process works because selection favors those genotypes that respond most readily to the stimulus.

The ‘‘Baldwin effect", the ‘‘stabilizing selection" of Schamlhausen and the ‘‘genetic assimilation" of Waddington are Darwinian processes that stress the active role that phenotypes have in the evolutionary process. Learning and evolution are, within this line of thought, cause and consequence of themselves.

In the next section we review some models that provide support of the existence of a strong interaction between learning and biological evolution.

3. Genetic Assimilation

In a classical experiment that has been confirmed by different scientists, Waddington acted upon larvae of the fruit fly *D. melanogaster* with an external environmental stimulus that consisted in a heat shock during the larvae development. This makes disappear the a common character consisting in the presence of a transversal vein in the wings. Waddington named the corresponding phenotype *crossveinless* wings or **CVL** for short. Geneticists call this characters ‘‘phenocopies" because are phenotypic departures that resemble true genetic mutations.

The experiment consisted in producing such phenocopies by a heat shocks and then only allow to reproduce the **CVL** phenotypes. By performing this artificial selection process, the percentage of **CVL** phenotypes progressively increase through successive generations. This fact is nothing strange in itself because confirmed that the **CVL** phenotype is associated to some kind of genetic information that could be selected upon. The surprising result of the experiment was to check that after the 14-th generation, all descendants started to show the **CVL** trait regardless of the fact that they where not submitted to heat shock. In this way a feature that was induced by an action during the development of the animal in early generations ended by being fixed in the genotype of the later ones. According to the words of Waddington a ‘‘canalization" has been produced that favored the genetic assimilation of the **CVL** phenotype in the whole population.

E. Jablonka and M. Lamb (1995) made a recent neo-Darwinian interpretation of the process of genetic assimilation. Assume that in order to encode the **CVL** phenotype three genetic loci *A*, *B* and *C* are involved, each one with two allelic variations that we denote A_1 and A_2; B_1 and B_2 and C_1 and C_2. There are therefore 27 possible genotypes (see Table 1). In genetic language this is said to be a ‘‘polygenic trait". The genes involvedin such trait interact through epistatic effects in order to express themselves as the **CVL** phenotype.

Let us consider that in normal conditions of development *envr.1* all possible combinations produce a normal phenotype (**N**) with the crossed vein except for one that produces the **CVL** phenotype in a constitutive manner (**CON**). Assume next that the expression of the **CVL** phenotype is induced (**IND**) through a heat shock *envr.2* the expression of the phenotype **N**, **CON** or **IND** are governed by the following four rules:

1. $A_1 A_1 B_2 B_2 C_2 C_2$ is **CON** in *envr.1* and *envr.2*
2. All genotypes $A_1\, X\, B_2\, X\, C_2\, X$ with the only exception of $A_1\, A_1\, B_2\, B_2\, C_2\, C_2$ are *IND* in *envr.2*
3. The genotypes $A_2\, X\, B_1\, X\, C_2 C_2$ are IND in *envr.2*
4. The combinations involving $C_1 C_1$ and those not mentioned in statements 1 and 3 are **N** in both environments

With these rules the probability that a population contains the genetic combination that expresses the **CVL** phenotype in a constitutive manner (**CON**) depends upon the frequency of the alleles and of the number of flies in the population. If the initial frequencies are: A_1 = 0.5 ; A_2 = 0.5; B_1 = 0.9 ; B_2 = 0.1; C_1 = 0.8 ; C_2 = 0.2. The probability to obtain the genotype $A_1\, A_1\, B_2\, B_2\, C_2\, C_2$ is $0.5^2 x 0.1^2 x 0.2^2 = 0.0001$. That is to say that one out of 10^4 flies show the **CON CVL** phenotype. In spite of the fact that Waddington did not start his experiment with such a high number, the application of the environmental stimulus and the iterated selection of **IND** and **CON** phenotypes in successive generations forced an increase in the frequency of the alleles A_1, B_2 and C_2 at the expense of A_2, B_1 and C_1. After many generations of artificial selection a high proportion of the population was formed by genotypes $A_1\, A_1\, B_2\, B_2\, C_2\, C_2$ and therefore there was no need of the heat shock because the **CVL** phenotype was already **CON**.

The process works because selection favors those genotypes that produce a response that is better tuned to the environmental stimulus, i.e. (artificial) selection favors a change in the norm of reaction exactly as suggested by Schmalhausen. The environment in this case played a role that usually has not been recognized in the general models of population genetics. From the classic point of view and in agreement with the model of allelic substitutions, selection acted upon existing genetic variations that end up by increasing in the population helped by environmental conditions.

The process of genetic assimilation has however a drastic limitation. If the trait does not exists, the population will have to wait for a purely random, beneficial mutation that could give origin to it. In Waddington’s experiment albeit the **CVL** mutant did not exist originally in the population and probably it did not appear during the experiment as a consequence of a random mutation, the evolutionary change took place very rapidly. Waddington insisted that genetic assimilation during evolution was an important effect because it provided a mechanism by which population could give a rapid adaptive answer to the change of the environment.

Table 1. Genotypes and possible phenotypes.

Genotypes	Phen. envr. 1	Phen. envr. 2
$A_1 A_1 B_1 B_1 C_1 C_1$	N	N
$A_1 A_1 B_1 B_1 C_1 C_2$	N	N
$A_1 A_1 B_1 B_1 C_2 C_2$	N	N
$A_1 A_1 B_1 B_2 C_1 C_1$	N	N
$A_1 A_1 B_1 B_2 C_1 C_2$	N	**IND**
$A_1 A_1 B_1 B_2 C_2 C_2$	N	**IND**
$A_1 A_1 B_2 B_2 C_1 C_1$	N	N
$A_1 A_1 B_2 B_2 C_1 C_2$	N	**IND**
$A_1 A_1 B_2 B_2 C_2 C_2$	CON	**CON**
$A_1 A_2 B_1 B_1 C_1 C_1$	N	N
$A_1 A_2 B_1 B_1 C_1 C_2$	N	N
$\mathrm{A_1 A_2 B_1 B_1 C_2 C_2}$	N	**IND**
$\mathrm{A_1 A_2 B_1 B_2 C_1 C_1}$	N	N
$\mathrm{A_1 A_2 B_1 B_2 C_1 C_2}$	N	**IND**
$\mathrm{A_1 A_2 B_1 B_2 C_2 C_2}$	N	**IND**
$\mathrm{A_1 A_2 B_2 B_2 C_1 C_1}$	N	N
$\mathrm{A_1 A_2 B_2 B_2 C_1 C_2}$	N	**IND**
$\mathrm{A_1 A_2 B_2 B_2 C_2 C_2}$	N	**IND**
$\mathrm{A_2 A_2 B_1 B_1 C_1 C_1}$	N	N
$\mathrm{A_2 A_2 B_1 B_1 C_1 C_2}$	N	N
$\mathrm{A_2 A_2 B_1 B_1 C_2 C_2}$	N	**IND**
$\mathrm{A_2 A_2 B_1 B_2 C_1 C_1}$	N	N
$\mathrm{A_2 A_2 B_1 B_2 C_1 C_2}$	N	N
$\mathrm{A_2 A_2 B_1 B_2 C_2 C_2}$	N	**IND**
$\mathrm{A_2 A_2 B_2 B_2 C_1 C_1}$	N	N
$\mathrm{A_2 A_2 B_2 B_2 C_1 C_2}$	N	N
$\mathrm{A_2 A_2 B_2 B_2 C_2 C_2}$	N	N

4. The Baldwin Effect

In order to discuss the Baldwin effect on a quantitative basis, G. Hinton and S. Nowlan (1987) elaborated a theoretical model based in the framework of artificial neural networks. The model was geared to investigate how learning could influence evolutionary changes. The model assumes a living being that has a hypothetical neural network involving, say, 20 synaptic connections that are either excitatory (labeled by a **1**) or inhibitory (labeled by a **0**). It is assumed that the animal having a single wrong connection in the network has the least possible reproductive success, only contributing to the population of next generation with its survival. Only the one that is born with the correct genetic information out of the set of $2^{20} \sim 10^6$ possibilities can leave a significant number of descendants.

This is equivalent to think that in a space of possible genotypes only one has the maximum possible reproductive success, and, at the same time, all the rest are equally inefficient. The **fitness landscape** looks like a pole planted in an otherwise flat surface. Under these circumstances the network could never evolve through the selection of random mutations because there isn't any gradual slope leading to the optimum, along which the population could be dragged by incremental changes. The situation is even more dramatic because any mutant appearing with the right synaptic connections as a consequence of purely random circumstances will have an ephimeral success if it has to reproduce sexually.

When this is the case, its genotype will rapidly disappear at the very moment of the reproduction because sex is tantamount to the exchange of genetic information between the parents.

The situation changes drastically if it is assumed that a third allele (labeled **?**) may be present, that expresses itself through a flexible feature in the phenotype. In this case, the neural network actually used by each individual could change due to the experience gathered during its lifetime. Such experiences could be considered to be the success or failure in the capture of a prey or the successful handling of a novel situation. Throughout the learning process, the result of each action (learning trial) is assessed.

To achieve success is the signature that the right combination of synapses has been found and learning stops. Furthermore, it is assumed that the individual that is able to find it in less trials, leaves a greater number of descendants to the next generation. The reproductive success of an individual is therefore directly related to the number of alleles that are equal to the optimum.

Descendants do not inherit the correct combination of synapses but rather the specific number and location of the flexible synapses of their ancestors. At difference with the case in which there is no learning, there is a gradual change in fitness as the optimum is approached. Individuals with less flexible connections and more correct alleles have a greater advantage over those that have too many **?** and less correct synapses. This is so because they require on the average less trials to find the optimum.

Let us assume for the sake of concreteness that half of the synaptic connections are fixed and the other that half are **?**. Furthermore assume that the search of the optimum combination is done by trial an error changing at random the **?** into **0**'s and **1**'s. Within a population such as this, there exists one individual that has the correct combination of fixed synapses among $2^{10} \sim 1000$. This sole individual can find the correct combination if it performs approximately 1000 learning trials. If it gets through, success is insured. The process has in addition a positive feedback, because any mutant descendant that changes flexible connections by correct fixed synapses will benefit from a shorter learning program because it will have less chance of doing mistakes during that period. In this way natural selection will produce a greater number of individuals with innate abilities.

In theory, the correct connection will not get fixed completely because the selection pressure gets weaker as the population tends to be more homogeneous and with similar innate features. In this way, learning eases the process of development of innate characters largely accelerating it with respect to the absence of genetic plasticity.

In fact, trying to find the right combination of correct synapses without the hints provided by the learning process is as trying to find a needle in a haystack. Learning is as having a metal detector to help in the search giving some information on how far away the needle is hidden (see discussion of Maynard Smith, 1987)

5. Hindrance of the Baldwin Effect

The above model has been improved recently by adding two ingredients geared to turn it more realistic (Dopazo, et al. 2001). On the one hand a greater ‘‘distance" is introduced between genetic information and the individual behavior and on the other individuals are endowed with different learning proficiencies.

This ansatz has been explored with mathematical models suitable to deal with infinite or finite populations. In the latter case, evolution is simulated numerically with the help of genetic algorithms. In the same fashion as explained for the model of Hinton and Nowlan, all individuals have a genetic information that involves three alleles **0**, **1** and **?** that express themselves respectively in fixed or adaptable synaptic connections of a neural network. This is taken to be a **perceptron** that involves as many inputs as alleles and provides a single output that is either a 1 or a 0. These two possible outputs represent the classification of data fed to the network. The output is considered to control the behavior of each individual. This situation can be assimilated to the case in which an edible object, say for instance a fruit, is presented to each individual in the population. This has to be classified into edible (output = 1) or not (output = 0) thus triggering two possible **behavioral responses.** This classification is performed processing input data (smell, color, size, texture, etc) that act as **environmental stimuli**.

Learning takes place because each member of the population is presented to M examples (i.e. sets of inputs) and flexible synapses have to be adapted in order to classify them properly. The behavioral response of each individual is assessed whether it is correct or not by comparing it with that provided by an optimal (reference) perceptron providing the correct responses. Environment is assumed to remain constant, thus the reference perceptron is assumed to remain unchanged during the whole evolutionary process. If the classification performed by an individual coincides with the one taken as a reference, a point is added to its possibilities of leaving descendants. Otherwise an adaptation (learning) algorithm is triggered that modifies the flexible synaptic connections to improve the wrong classification for future use with the remaining examples.

The genotypes in the population are originally determined randomly and modified in successive generations through selection and mutation. Having too many flexible alleles has a greater cost because the individual can make too many mistakes before a good learning is achieved. On the other hand an individual with many correct alleles has a very good classification performance. It is therefore clear that selection will favor the elimination of flexible alleles by replacing them with correct ones thus assimilating the environmental information of the reference perceptron through Baldwin effect.

The model has however also room for the assimilation of environmental information **without** learning. This can be achieved because a neural network having fixed synapses that differ from the optimum may nevertheless succeed in providing a good classification of a reduced number of examples thus gaining points towards its reproductive success. This feature is precisely the one that puts a greater distance between the genotype and the behavioral response thus enhancing the role played by the phenotype in the evolutionary process.

The combined effect of learning on the one hand, and the epistatic effects of a genotype without plasticity on the other, causes that the evolutionary process takes place in an effective space of allelic frequencies that is larger than the one studied by Hinton and Nowlan. In fact this is essentially restricted to the subspace of all genetic sequences that involve only correct and flexible alleles while in the perceptron model wrong synapses may also count for the reproductive success. In this extended model even in the case in which the **?** alleles are absent, evolution is possible and the optimal configuration is always reached. A comparison is thus possible between different ‘‘evolutionary paths" leading to the optimal configuration equal to that of the reference perceptron.

The numerical simulation confirms a general pattern that agrees with Baldwin’s suggestions: first wrong alleles are replaced by a combination of correct and flexible alleles. In a second stage, flexible alleles are replaced by correct ones. This path can in addition be compared to one in which individuals are constrained to have no flexible alleles. Evolution in this case is also successful in transcribing environmental information into the genotypes of the population. If analyzed in detail, the results are however somewhat surprising. After the same number of generations a population of rigid perceptrons **gets much closer** to the optimum than one having individuals with flexible alleles in an almost opposed picture as the one thought by Baldwin. However as far as reproductive success is concerned, the population of individuals that are capable of learning outperforms that of rigid perceptrons.

This pattern can also be studied allowing different learning proficiencies. This can be done by changing the number of examples M that are used during the training stage, a larger M corresponding to a higher learning proficiency. As M is increased the fastest genetic assimilation is achieved by the rigid perceptrons while a population capable of some learning stagnates away from the optimum. In the case of an infinitely effective learning capability evolution displays full stagnation far from the optimum: all flexible alleles are kept, all wrong alleles are eliminated but the number of correct alleles is not increased.

This can easily be understood. When learning is too effective there will be no selection pressure to increase the number of correct alleles at the expense of eliminating the flexible ones. As the cost (measured in number of descendants) of either learning or flexibility increases, Baldwin effect will become more effective. However if the cost is too high, the Baldwin effect will again be hindered in favor of the process of pure random allelic substitution.

6. Conclusions

Within the framework of the Evolutionary Synthesis, evolution is completely equivalent to the change of the allelic frequencies of the population. The classic model underlying this picture is one of continuous allelic substitution through random mutations and subsequent selection. With a much less frequency models have been thought in terms of the evolution of genetic plasticity with rather wide norms of reaction in order to adapt to changing environments. Equally poor has been the fortune of models based on the Baldwin effect that have been disqualified as being Lamarckian.

These alternative models of phenotypic evolution can today be understood within a purely Darwinian framework as a the result of the mutual influence of allelic substitution and phenotypic plasticity. When the costs of keeping phenotypic plasticity are greater than the benefits of adapting to a changing environment, the transfer of environmental information to the genotype will take place through random allelic substitutions. In addition, in the case in which environmental changes are very slow such transfer will almost be complete. The morphological or behavioral characters that were tested by the individuals of prior generations will thus be transferred into genetic information.

On the other hand when such costs are low, other evolutionary processes come into play in which plasticity plays a major role. The flexibility to develop new behavioral features through learning is, among all forms of organic adaptation, the one that can more rapidly produce drastic evolutionary changes. This is so because individuals deploy through their phenotypic plasticity, a wider set of potential organic adaptations, amplifying the sampling range of natural selection. The ability of learning and the later inheritance of acquired behavioral traits may well constitute one of the most effective ways of enhancing natural adaptation.

An acquired behavior can be changed into an innate disposition precisely due to fact that learning has a non vanishing cost. If adaptation is 100% effective there is no selection pressure to encode any information into the genotype. However, in general any living being who undergoes a learning process bears a cost. There are costs due to the time involved in the learning stage; due to mistakes done during that stage, or simply because there are organic limitations until the moment of reproductive maturity. When these costs are moderate, an accelerated transfer of information takes place from the environment to the genotype while this is hindered if learning is too effective.

E. Mayr, that has been one of the founders of the Evolutionary Synthesis suggested that for any species, during the invasion of a new ecological niche, behavioral adaptations could preceed the morphological ones (Mayr, 1963). What he did not consider is that the former could well help guiding the latter. In the context of linguistics, S. Pinker (1994) has suggested that the evolution of language is the ideal arena to test the relevance of the Baldwin effect. New verbal skills are continuously learned but it is also true that many of these are genetically assimilated. In the same fashion the anthropologist T. Deacon (1997) has insisted that the structural and functional innovations of the human brain are a consequence of a long co-evolution with linguistic abilities which is mediated principally by the Baldwin effect.

From the perspective of a general description of complex systems, learning and biological evolution can be regarded as being the same process but taking place in two different hierarchical levels. Ech of these are in turn associated to two widely different timescales. Learning can well be pictured as an evolutionary process that takes place during ontogeny. Through this process unsuccessful behaviors are progressively discarded in favor of those that are more beneficial for survival. The timescale of this process is of the order of the lifetime of the individuals. Biological evolution through allelic substitution cause that traits that pay less in reproductive success are discarded in favor of others that offer the possibility of a greater number of descendants. The timescale of this process has to be measured in generations and results in the adaptation or organisms to the (perhaps changing) surrounding environment.The results of experiments, observations and models that have been discussed above put the accent in the mutual influence of these two levels of organization across widely different timescales. Within this picture, the Baldwin effect can well be regarded as the result of the mutual influence of two evolutionary processes pertaining to two different hierarchical level of organization of life systems.

Acknowledgements

HD wishes to acknowledge the support of CONICET through a postdoctoral fellowship RP acknowledge economic support from the EU- research contract ARG/B7-3011/94/97. HD and RP hold a UBA-research contract UBACYT-PS021, 1998/2001.

References

Baldwin, J. M. (1896). A new factor in evolution. *American Naturalist.* 30, 441-451.

Bellew, R. & Mitchell M. (1996). *Adaptive Individuals in Evolving Populations: models and algorithms.* Vol XXV B, Santa Fe Institute. Addison Wesley Publ. Co Inc NY.

Deacon, T. (1992). *The Symbolic Species: the coevolution of language and the brain.* W. Norton Publ. Co. NY.

Dopazo H., Gordon M. B., Perazzo R. & Risau-Gusman S. (2001). A model for the interaction of learning and evolution. *Bull of Math. Biol.* 63, 117-134.

Hinton, G. E. & Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems* 1, 495-502.

Jablonka, E. & Lamb, M. 1994. *Epigenetic Inheritance and Evolution: The Lamrkian Dimension.* Oxford University Press. Oxford.

Maynard Smith, J. 1987. When learning guides evolution. *Nature* 349, 761-762.

Mayr, E. (1963). *Animal Species and Evolution.* Cambridge. Massachusetts: Harvad Univ. Press.

Morgan, C.L. (1896). On modification and variation. *Science* 4, 733-740.

Osborn, H. F. (1896). Ontogenic and phylogenic variation. *Science* 4, 786-789.

Schlichting, C. & Pigliucci, M. (1998). *Phenotypic Evolution: a Reaction Norm Perspective.* Ass. Inc. Publishers. Sunderland, Massacussetts.

Simpson, G.C. (1953). The Baldwin effect. *Evolution* 7, 110-117.

Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characters. *Nature* 150, 563-565.

Other publications concerning the Baldwin effect may be consulted in: http://www.ai.mit.edu/~joanna/baldwin.html

